2015

Insights into Chlorinated Aliphatic Alkane Reductive Dehalogenation by Dehalogenimonas Lykanthroporepellens BL-DC-9T and Development of Nucleic Acid-Based Techniques for Monitoring Gene Expression

Kalpataru Mukherjee
Louisiana State University and Agricultural and Mechanical College

INSIGHTS INTO CHLORINATED ALIPHATIC ALKANE REDUCTIVE DEHALOGENATION BY *DEHALOGENIMONAS LYKANTHROPOREPELLENS* BL-DC-9^T AND DEVELOPMENT OF NUCLEIC ACID-BASED TECHNIQUES FOR MONITORING GENE EXPRESSION

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Biological Sciences

by
Kalpataru Mukherjee
B.S., Bangalore University, 2005
M.S., VIT University, 2007
August 2015

Acknowledgements

I would like to take this opportunity to thank and acknowledge my advisor Dr. Fred A. Rainey for offering me the opportunity to pursue my Ph.D. research at the department of Biological Sciences at Louisiana State University.

I am sincerely thankful to my illustrious co-advisor Dr. William M. Moe for his insightful guidance throughout my Ph.D. research. It has been a great learning experience for me to work in his research group and his continuous support has been a great motivation to complete my research and lay the foundation stone for a scientific career.

I thank my executive committee member Dr. Anne Grove for her valuable suggestions and I acknowledge that her ideas helped me to troubleshoot a lot of critical problems in my research. I am also thankful to my executive committee member Dr. Grover L. Waldrop.

I thank the members of GenLab at LSU school of Veterinary medicine and the Genomics facility centreat Biological sciences LSU for running the DNA sequencing reactions and the chip assays for my RNA work. I am thankful to the past and present members of our research group Dr. Kimberley Bowman, Dr. Jyoti U Rao who helped me immensely during my initial days at LSU and also Trent A Key, Hector Narez, Jacob Dellehay, and Jie Chin for their friendship and support.

I have enjoyed great suppport and company of my friends Dr. Anirban Sarkar, Agnimitro Chakrabarti, Girish Yajurvedi, Lokendra Thakur, Subhajit Siddhanta, Rakesh Gunaseelan, Dr. Janki Chadderwala, Vikram Sadhya, Tapan Sarkar, Sonhita Sarkar and Unmesh Krittikar and these people have made me feel at home in spite of being in a foreign country.

Lastly, I profoundly appreciate the love, support, encouragement and sacrifices of my family members Rituparna Senapati, Chandan Mukherjee and Sumita Mukherjee and my brother Mainak Mukherjee for their consistent and wholehearted love and support.

Table of Contents

List of Tables

List of Figures

Abstract

Dehalogenimonas lykanthroporepellens BL-DC-9^T is an anaerobic bacterial strain that can reductively dehalogenate a variety of vicinally chlorinated alkanes. Making use of the strain's genome sequence, PCR primers were designed to uniquely target each of the 25 reductive dehalogenase homologous (*rdhA*) genes present in the genome. RNA extracted from *D. lykanthroporepellens* BL-DC-9^T cultures that were dechlorinating three chlorinated solvents (1,2-dichloroethane, 1,2-dichloropropane and 1,2,3-trichloropropane) were used in reverse transcriptase PCR (RT-PCR) to determine *rdhA* gene expression patterns. Transcripts from 19 *rdhA* genes were detected in the RT-PCR experiments, with identical gene expression patterns during growth with all three chlorinated electron acceptors employed in the study. The transcripts included thirteen full-length *rdhA* genes, six truncated *rdhA* genes, and five *rdhA* genes having a cognate *rdhB* gene. Transcripts from four putative transcriptional regulators were also detected during organohalide respiration with 1,2-dichloroethane, 1,2-dichloropropane and 1,2,3-trichloropropane, providing insights in the regulatory controls on *rdhA* gene expression.

A dehalogenase enzyme assay was developed, and attempts were made to heterologously express the *rdhA* gene with locus tag Dehly_1524 (recently identified as encoding a 1,2-dichloropropane reductive dehalogenase) in *Escherichia coli*. The majority of the heterologously expressed protein was found in the insoluble fraction and in a catalytically inactive form under all conditions tested; this was inferred to be due to the absence of *de novo* cobamide synthesis in *E. coli* and the requirement for this cofactor for enzymatic activity.

Functional expression of the Dehly_1524 protein was also attempted under anoxic conditions using the cobamide producing species *Shimwellia blattae* that was recently identified

as having a genetic system that allowed the functional expression of dehalogenases from another species (*Desulfitobacterium hafniense* strains Y51 and DCB-2). The Dehly_1524 protein was not catalytically active when expressed in *Shimwellia blattae* under the conditions tested, however, indicating that *Shimwellia blattae* may not contain a genetic system broadly suitable for functional expression of the growing number of putative reductive dehalogenase genes available in the public databases.

Chapter 1. Introduction and literature review

1.1 Introduction

Dehalogenimonas lykanthroporepellens strain BL-DC-9^T was first isolated from contaminated groundwater at the PetroProcessors of Louisiana, Inc. Superfund site located near Baton Rouge, LA (USA). This species is of particular interest because of its unique ability to reductively dehalogenate a variety of vicinally chlorinated aliphatic alkanes including 1,2-dichloroethane (1,2-DCA), 1,2-dichloropropane (1,2-DCP), 1,1,2,2-tetrachloroethane (1,1,2,2-TeCA), 1,1,2-trichloroethane (1,1,2-TCA), and 1,2,3-trichloropropane (1,2,3-TCP) via dihaloelimination reactions (Moe *et al*., 2009; Yan *et al*., 2009). The recently sequenced genome of *D. lykanthroporepellens* strain BL-DC-9^T (GenBank accession no. CP002084) contains 25 loci having reductive dehalogenase homologous (*rdh*) genes scattered throughout the chromosome (Siddaramappa *et al*., 2012). Prior to the research presented in this dissertation, the expression pattern of these *rdh* genes during reductive dechlorination was unknown. Research presented in this dissertation was aimed at determining the pattern of *rdh* gene expression during dechlorination of three substrates: 1,2-dichloroethane (1,2-DCA), 1,2-dichloropropane (1,2-DCP), and 1,2,3-trichloropropane (1,2,3-TCP). Efforts were also made to heterologously express a functional reductive dehalogenase enzyme using two different bacterial hosts.

1.2 Dissertation Organization

Chapter 1 of this dissertation presents a literature review focused on the characteristics of dehalogenase enzymes, previous transcriptomic studies of dehalogenating bacteria, and heterologous expression of dehalogenase enzymes. Chapter 2 describes experiments that employed reverse-transcriptase PCR (RT-PCR) to study the pattern of reductive dehalogenase

gene expression during the dechlorination 1,2-dichloroethane (1,2-DCA), 1,2-dichloropropane (1,2-DCP), and 1,2,3-trichloropropane (1,2,3-TCP). It also includes results from experiments conducted to determine gene expression patterns of the transcription factors that are associated with a subset of reductive dehalogenase genes in *Dehalogenimonas lykanthroporepellens* strain BL-DC-9^T. Chapter 3 presents results from experiments aimed at expression of a functional reductive dehalogenase protein in *Escherichia coli*. Chapter 4 presents the results from experiments aimed at functional expression of the reductive dehalogenase protein in a cobamide producing *Shimwellia blattae* strain. Chapter 5 presents a summary of the major findings from the research and recommendations for future work.

1.3 History and overview

Reductive dehalogenation is a bioremediation process in which microbes capable of transforming toxic xenobiotic compounds are exploited for biodegradation of environmental pollutants and industrial wastes (Holliger *et al*., 1999). Anaerobic bacterial strains have been identified that use the polychlorinated solvents as electron acceptors and derive energy from reductive reactions. This process in which anaerobic bacterial strains derive energy from this process has been alternatively referred to in the literature as dehalorespiration or organohalide respiration. The enzymes that bring about the anaerobic transformation are referred to as reductive dehalogenases (Holliger *et al*., 2003).

Polychlorinated alkanes are some of the most common groundwater contaminants in the United States and throughout the world (Löffler and Edwards, 2006). Many are known or suspected to cause adverse health effects, with several belonging to the category of suspected carcinogens (De Wildeman and Verstraete, 2003).

D. lykanthroporepellens strain BL-DC-9^T has been reported to dechlorinate a number of polychlorinated alkanes. It dechlorinates 1,2-dichloroethane to the non-toxic, non-halogenated final product ethene. Similarly, it has been shown to dechlorinate 1,2-dichloropropane to the non-chlorinated, non-toxic product propene. It is also one of the few strains isolated to date that can anaerobically dechlorinate 1,2,3-trichloropropane, forming allyl chloride (3-chloro-1-propene), which is relatively unstable and can undergo a variety of abiotic reactions to produce a variety of products includig allyl alcohol (Moe *et al*., 2009; Yan *et al*., 2009).

Tetrachloroethene (PCE) has been found to be dechlorinated to trichloroethene (TCE) by bacterial strains having the *pceA* gene, which codes for proteins capable of dechlorination of PCE. The ε-proteobacterium *Sulfurospirillum multivorans* strain K performs the dechlorination of PCE (Neumann *et al*., 1998). *Dehalobacter restrictus* strain PER-K23, *Desulfitobacterium hafniense* strain TCE1, *D. hafniense* strain PCE-S and *Desulfitobacterium* sp. Strain Y51 are also reported to have the *pceA* gene, which codes for the protein capable of dechlorinating PCE (Maillard *et al*., 2003). The *pceA* gene has also been reported in *Dehalococcoides mccartyi* strain 195 (Magnuson *et al*., 2000).

Representatives of the species *D. mccartyi* are able to dechlorinate a variety of halogenated pollutants including chlorobenzenes and chlorinated ethenes. Of particular significance in the environmental field, representatives of this species are the only strains thus far isolated in pure culture with an ability to anaerobically dehalogenate the carcinogen vinyl chloride (Löffler *et al*., 2013). Interestingly, while *D. mccartyi* strains appear to be obligate organohalide respirers not known to grow using electron acceptors other than halogenated organics, the various isolates obtained in pure culture to date vary with respect to which chlorinated compounds they dehalogenate. For example, *D. mccartyi* strain CBDB1

dechlorinates chlorobenzenes and dibenzodioxin congeners. *D. mccartyi* strains 195 and FL2 dechlorinates polychlorinated ethenes, and *D. mccartyi* strain BAV1 is capable of dehalogenating DCE and vinyl chloride (Cupples *et al*., 2003; He *et al*., 2003; Maymó-Gatell *et al*., 1997, Maymó-Gatell *et al*., 2001).

Presently, there are more than 650 reductive dehalogenase homologous (*rdh*) gene sequences available in the public databases. The functions of only a few of them, however, are known (Tang *et al*., 2012). Extreme oxygen sensitivity and slow growth of the organisms make the characterization of these genes and the proteins that they encode difficult. Although a genetic system for expression of functional reductive dehalogenase has recently been identified, production of functional reductive dehalogenases still remains challenging (Mac Nelly *et al*., 2014).

1.4 Reductive dehalogenase enzyme characteristics

At the genetic level, reductive dehalogenase proteins have been classified into three types. Most of the proteins characterized to date have the unique similarity of possessing two iron-sulfur clusters along with a corrinoid cofactor (Holliger *et al*., 2003). At the amino-terminal end, the proteins are characterized by a conserved twin-arginine transport system (TAT) signal sequence that is thought to aid in the export of periplasmic proteins (Berks *et al*., 2000). The second type of protein is characterized by the absence of iron sulfur cluster and the presence of one corrinoid cofactor. Examples for this class of proteins include the ortho-chlorophenol dehalogenase from *Desulfitobacterium frappieri* PCP-1 (Boyer, 2003) and the PCE reductive dehalogenase that was originally attributed to *Clostridium bifermentans* DPH-1 (Okeke, 2001), a culture that was later determined to include two populations, a *Clostridum* sp. and

tetrachloroethene-dechlorinating *Desulfitobacterium hafniense* strain JH1 (Fletcher *et al*. 2008). The third type of protein is totally different from the first two types and is a heme protein having two subunits. This protein is unique to the chlorobenzoate dehalogenase of *Desulfomonile tiedjei* (Ni *et al*., 1995). For the small number of reductive dehalogenases that have been functionally characterized to date, the genes are organized in the form of an operon. For example, the genes encoding tetrachloroethene reductive dehalogenase (*pceA*) and trichloroethene reductive dehalogenase (*tceA*) are closely linked to genes *pceB* and *tceB*. These latter genes (i.e., *pceB* and *tceB*) are thought to code for hydrophobic proteins that acts as membrane anchors for the reductive dehalogenases and are transcribed along with the first part (i.e., *pceA* and *pceB* are co-transcribed) (Villemur *et al*., 2002).

If a "full-length" *rdhA* is predicted to encode a protein containing a twin-arginine sequence at the N-terminus, two iron-sulfur cluster binding motifs at the C-terminus, and an intervening sequence of ~450-500 amino acids, then *D. lykanthroporepellens* BL-DC-9^T contains 17 such genes (Siddaramappa *et al*., 2012). Eight other *rdhA* genes in the genome appear to be substantially truncated and are predicted to encode proteins lacking the N-terminus, C-terminus, or both (Table S1). Only six of the *rdhA* ORFs have a cognate *rdhB*, and an additional *rdhB* gene (Dehly_1504) appears to be an orphan with no cognate *rdhA* ORF nearby (Siddaramappa *et al*., 2012)

1.5 [Fe-S] Clusters functional significance

As mentioned earlier in this chapter, reductive dehalogenase enzymes belong to a class of enzymes that have iron sulfur clusters. Iron-sulfur cluster binding proteins are a family of proteins that have Fe and S as cofactors and are involved in electron transfer pathways in redox

reactions (Rao *et al*., 2004). It is recurringly observed in [Fe-S] cluster binding proteins, that the iron element is liganded by a number of cysteine residues, and these families of proteins are widespread in a large number of biological samples (Gorst *et al*., 1995). They have also been detected in anaerobic bacteria through electron paramagnetic resonance signals (Shethna *et al*., 1964). Detection of electron paramagnetic signals is the most frequent way of confirming the presence of [Fe-S] clusters in proteins although a combination of multiple spectroscopic methods (e.g., UV spectroscopy) are generally the procedure for the detection of [Fe-S] clusters (Lill *et al*., 2006).

The [Fe-S] clusters function in mediating a single electron transfer process in bacterial cells. A specific example of an iron sulfur binding protein is the [3Fe-4S] bacterial enzyme Ferrodoxin I (Johnson *et al*., 2005). The [Fe-S] binding proteins mediate oxidation-reduction reactions by switching between the different oxidized and reduced forms. Rhombic [2Fe-2S], for example, can carry out the redox reactions by interchanging between the oxidized form $[2Fe\text{-}2S]^{2+}$ to the reduced form $[2Fe\text{-}2S]^{+}$ (Rao *et al*.,2004). In a similar manner, Cubane [4Fe-4S] mediates redox reactions by existing in three different oxidation states: $[4Fe\text{-}4S]^{+}$, $[4Fe\text{-}4S]^{2+}$, and $[4Fe\text{-}4S]^{3+}$ (Rao *et al*., 2004). Although a large number of [Fe-S] binding proteins have been identified, a clear binding motif for binding of [Fe-S] has yet to be identified. Some of the common structural residues for the identification of [Fe-S] cluster proteins are the residues $CX_4CX_2CX_{30}C$ in ferredoxins and the residue $CX_2CX_2CX_{20\text{-}40}C$ for [4Fe-4S] (Lill *et al*., 2006). Chemical reconstitution of a number of functionally inactive [Fe-S] binding proteins have resulted in catalytically active proteins, and several [Fe-S] biogenesis pathways and genes involved in [Fe-S] biogenesis has been identified in prokaryotes (Malkin *et al*., 1966, Barras *et al*., 2005).

1.6 Overview of catalysis by cobalamine dependent enzymes

1.6.1 Structure

Cobalamines are some of the chemical forms of Vitamin B_{12} and are found in different conformations in B_{12} dependent enzymes. Cobalamines found in B_{12} dependent enzymes contain a central cobalt atom that is liganded with four nitrogen atoms from the pyrroles of the corrin ring (Fig 1.1). The upper ligand attached to the cobalt atom is deoxyadenosine in the case of adenosine cobalamine, methyl in methylcobalamin, hydroxy in hydroxocobalamin and cyanocobalamin in vitamin B_{12} (Figure 1.1). The lower ligand attached to the cobalt varies for different organisms, and the most commonly occurring lower ligand for cobalamines is the endogenous base dimethylbenzimidazole (Fig 1.1). At physiological pH, the lower ligand generally exists as a base on confirmation; however, it can also exist as a base off conformation when the lower base is protonated. In some instances, the dimethylbenzimidazole can be replaced by a histidine residue from the protein, which attaches to the corrin ring (Brown *et al*., 1984; Mancia *et al*., 1986). Tight binding to the associated protein is achieved by hydrogen bond formations between the regulated protein and the peripheral side chains of the corrin ring (Banerjee and Ragsdale, 2003).

Vitamin B_{12} is a family of organometallic cofactors that are known to regulate different enzymes such as isomerases, methyltransferases, and more recently, dehalogenases. The role of vitamin B_{12} in the catalytic activity by isomerases and methyltransferases has been well studied, while their role in free radical reactions in dehalogenases is not yet completely elucidated (Banerjee and Ragsdale, 2003).

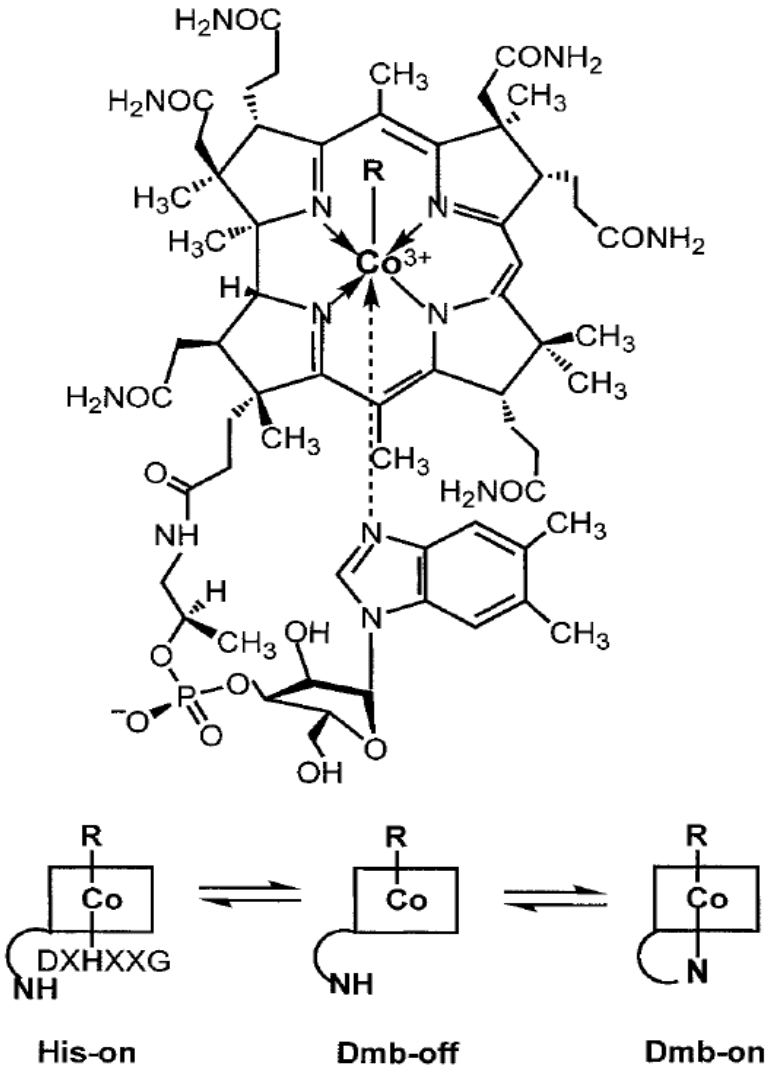


Figure 1.1 General structure of the cobalamine found in vitamin B_{12} dependent enzymes. R is the upper ligand attached to the cobalt atom (Reprinted with permission from Banerjee and Ragsdale, 2003). The upper ligand is deoxyadenosine in case of adenosinecobalamine, methyl in methylcobalamine, hydroxy in hydroxocobalamine and cyanocobalamine in vitamin B_{12}. The lower ligand dimethylbenzimidazole can exist as base on confirmation (Dmb-on), base off (Dmb-off, protonated form), or His on conformation. See appendix E2 for permission obtained for reprint.

1.6.2 Proposed reaction mechanisms for cobalamine dependent dehalogenases

For the cobalamin dependent dehalogenases, it has been assumed that the corrinoid is the site of reductive dehalogenation (Krone *et al*., 1989). The first proposed mechanistic model predicts that a reduction reaction of the organohalides takes place through a Co(I) free radical and no organocobalamin intermediate is formed. An alternate model predicts that an organocobalamin product is formed at the first step, and the halogen atom is eliminated in the subsequent step in which an aryl-Cobal (III) intermediate is formed (Banerjee and Ragsdale, 2003). As described in detail in Section 1.7, the recently determined crystal structures of the PceA reductive dehalogenase from *Dehalospirillium multivorans* (Bommer *et al*., 2014) and NpRdhA protein from *Nitratireductor pacif*icus pht -3B (Payne *et al*., 2015) has provided further insights into the free radical reaction pathway of the dehalogenase enzymes.

1.7 Structures and mechanistic insights into dechlorination pathways of functionally characterized enzymes

The first report of an X-ray crystal structure of a functional reductive dehalogenase enzyme was that of the tetrachloroethene reductive dehalogenase PceA from *Dehalospirillium multivorans* (Bommer *et al*., 2014). The crystal structure of this enzyme in the empty state and when bound to substrate provides insights into the cobalamine supported dehaloelimination reactions catalyzed by this family of enzymes.

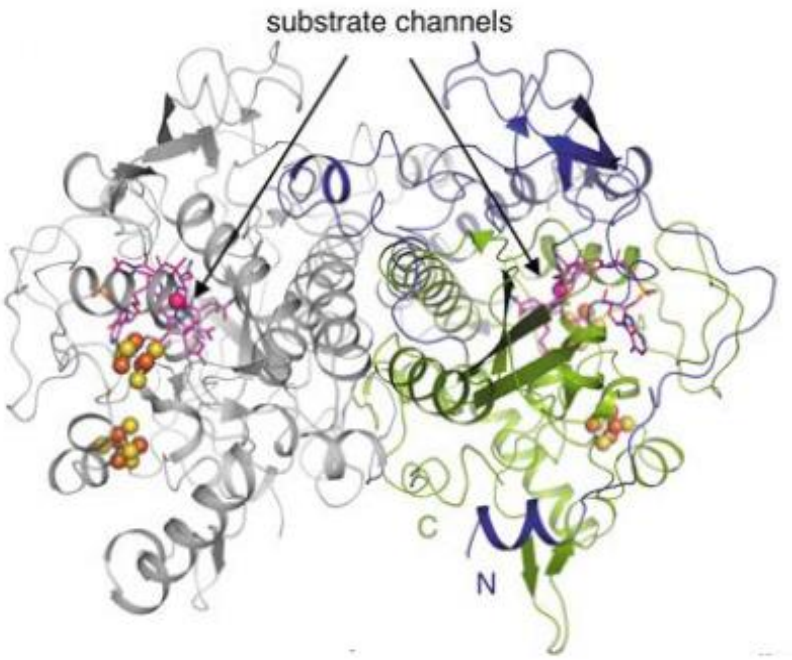


Figure 1.2 Crystal structure of the dimeric PceA reductive dehalogenase enzyme from *Dehalospirillium multivorans* (Bommer M, Kunze C, Fesseler J, Schubert T, Diekert G and Dobbek H (2014) Structural basis of organohalide respiration. *Science* **6208**: 455-8. Reprinted with permission from AAAS). One protomer is shown in grey while the other protomer is colored. The cofactor binding regions are shown in green color whereas the variable regions are in blue. The two [4Fe-4S] are shown as spheres and the corrinoid norpseudo vitamin B_{12} as purple coloured sticks.

The structure shows that the enzyme exists in the form of the dimer, and each protomer consists of two [4Fe-4S] clusters, a norpseudo vitamin B_{12} binding core and insertion units along with N-terminal and C terminal regions (Fig 1.2). The corrinoid norspeudo vitamin B_{12} has been characterized as the cofactor for this enzyme in the *Dehalospirillium multivorans* species (Krautler *et al*., 2003). The proximal and distal [4Fe-4S] clusters to the cobalt atom of the corrin ring are located at the enzyme surface whereas the cofactor binding site is located deep in the enzyme structure (Bommer *et al*., 2014). It can be assumed that swift electron transfer takes place inside the protein monomer as the two [4Fe-4S] clusters are located close to the cobalt atom of the corrin ring (Fig 1.2). The cofactor binding to the cofactor binding domain of the

enzyme is stabilized by hydrogen bonds between the norpseudovitamin B_{12} ring system and the protein matrix (Bommer *et al*., 2014). It can also be seen that there is one active site in each monomer, resulting in two active sites per dimer of the characterized enzyme.

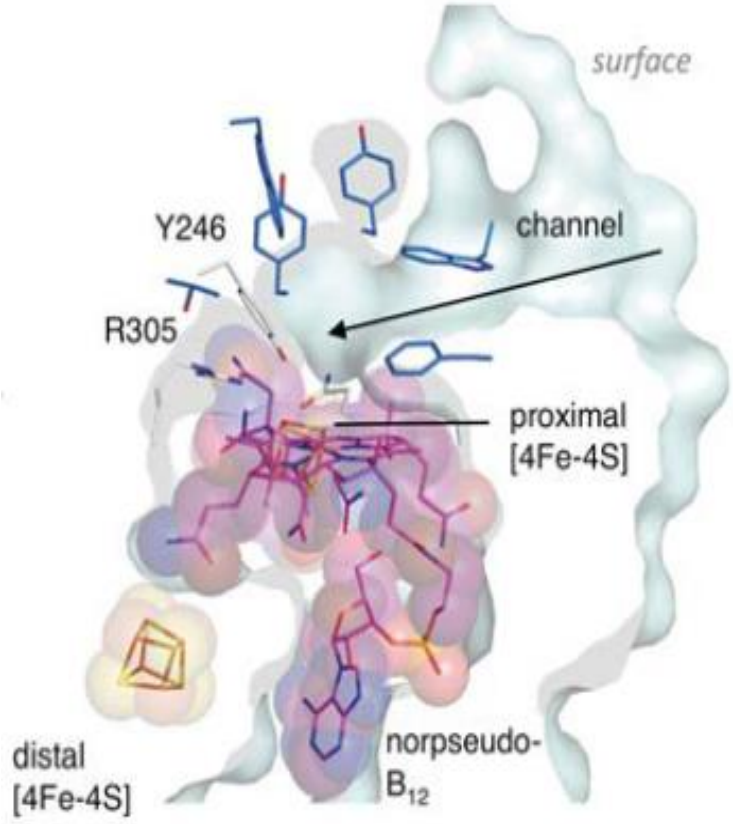


Figure 1.3 Crystal structure of the active site of the PceA reductive dehalogenase enzyme from *Dehalospirillium multivorans* (Bommer M, Kunze C, Fesseler J, Schubert T, Diekert G and Dobbek H (2014) Structural basis of organohalide respiration. *Science* **6208**: 455-8. Reprinted with permission from AAAS). The arrow shows the substrate entry channel into the active site of the protein. White and Blue sticks indicate the location of the polar amino acids and hydrophobic amino acids at the active site pocket.

For the chloroethene substrates to reach the active site of the protein they have to pass through a selection filter created by the residues from the N-terminal end of the protein (Bommer *et al*., 2014) and then though a 12 Å hydrophobic channel leading to the active site (Fig 1.3). The active site pocket of the PceA protein is bordered by Tyr246, Arg305 and Asn272 out of which

the Tyrosine residue has been found to be conserved among reductive dehalogenases and is positioned such that its side chains point to the active site of the enzyme (Fig 1.3 and Bommer *et al*., 2014). In addition to the substrate filter, the side chains of the amino acids lining the active site also function as a selection filter for substrate binding at the active site. The rigid conformations of the side chains allow binding of only specific substrates and hence permits binding of substrates that are only compatible in structure at the substrate binding site (Bommer *et al*., 2014).

The organohalide respiration mechanism involves the transfer of two electrons to the halogenated carbon substrate and transfer of a proton to the leaving chloride ion (Bommer *et al*., 2014). The first step in the mechanism of reductive dehalogenase catalysis in the PceA enzyme is likely to be the reduction of the neutral cobalt(II) in the PceA to the reactive cobalt(I) at the active site by weak ligand binding (Bommer *et al*., 2014). Hence the first electron transfer is possibly from the supernucleophile cobalt (I) to perchloroethene resulting in the formation of the trichlorovinyl radical (Shey *et al*., 2000). This PceA enzyme structure model proposes that the second electron transfer takes place from the [4Fe-4S] cluster. Transfer of a proton from the conserved tyrosine residue at the active site of the enzyme to stabilize the leaving chloride ion is perhaps the final step (Glod *et al*., 1997). Positively charged Arg305 at the active site pocket conceivably balances the proton removal from the tyrosine residue.

The second structure of a functional reductive dehalogenase that was reported was the crystal structure of the NpRdhA protein from *Nitratireductor pacificus* pht -3B (Payne *et al*., 2015).

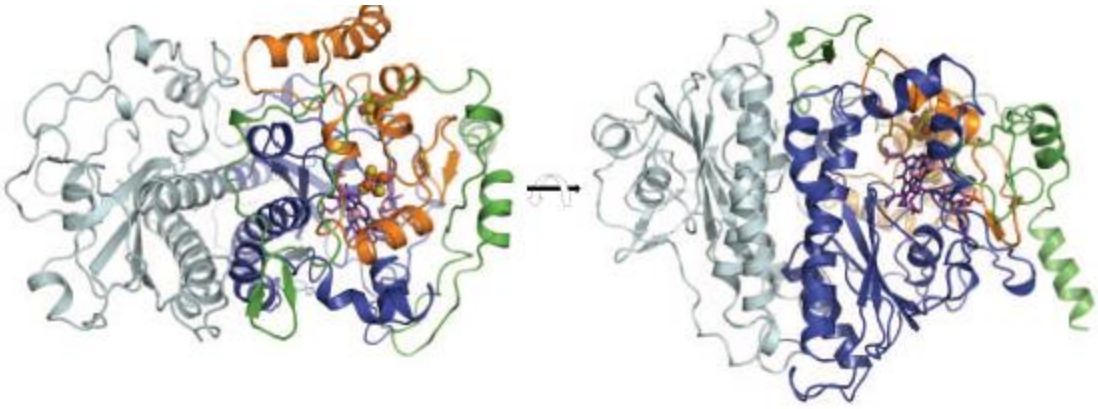


Figure 1.4 Crystal structure of the dimeric NpRdhA reductive dehalogenase enzyme NpRdhA protein from *Nitratireductor pacificus* pht -3B (Payne *et al*., 2015). The monomer is shown in two orientations The N-terminal domain is shown in grey. The B_{12} binding region is shown in blue. The Fe-S binding regions are shown as spheres and the flavin reductase type domains in green. Reprinted by permission from Macmillan Publishers Ltd: [Nature Publishing Group] (Payne KA, Quezada CP, Fisher K, Dunstan MS, Collins FA, Sjuts H, Levy C, Hay S, Rigby SEJ & Leys D (2015) Reductive dehalogenase structure suggests a mechanism for B12-dependent dehalogenation. *Nature* **7535**:513-516), copyright (2015). See appendix E2 for license terms from Nature Publishing Group.

The crystal structure of NpRdhA shows that the protein exists as a globular monomer comprising a central core domain for cobalamin binding along with N-terminal and C-terminal domains (Fig 1.4). The crystal structure also displays the existence of a flavin-reductase domain that aids in binding of the cobalamin cofactor in a base off form (Payne *et al*., 2015 and Fig 1.4). Similarity to the structure of the PceA protein exists in the fact that at the C-terminus of the protein, both the [Fe-S] binding clusters are located at the surface of the protein and [Fe-S] binding clusters are 9.8Å apart (Payne *et al*., 2015). The active site pocket of the NpRdhA protein has been recognized as the empty space above the cobalamin plane (Fig 1.5 Right). Docking of the protein with the substrate 3, 5-dibromo-4-hydroxybenzoic acid showed that the halogen, bromine atom lies in the active site of the enzyme and in close association with the

residues Tyr 426, Lys 488 and Arg 552 (Fig 1.5 Right). The docking also suggested that 3, 5-dibromo-4-hydroxybenzoic acid is bound to the active site in a deprotonated state as the hydroxyl groups of the substrate are hydrogen bonded to Ser 422, Lys 488 and Arg 552 at the active site pocket. The enzyme-substrate binding docking was further reinforced by the electron paramagnetic resonance spectroscopic study that showed an interaction between the halogen and cobalamin (Payne *et al*., 2015).

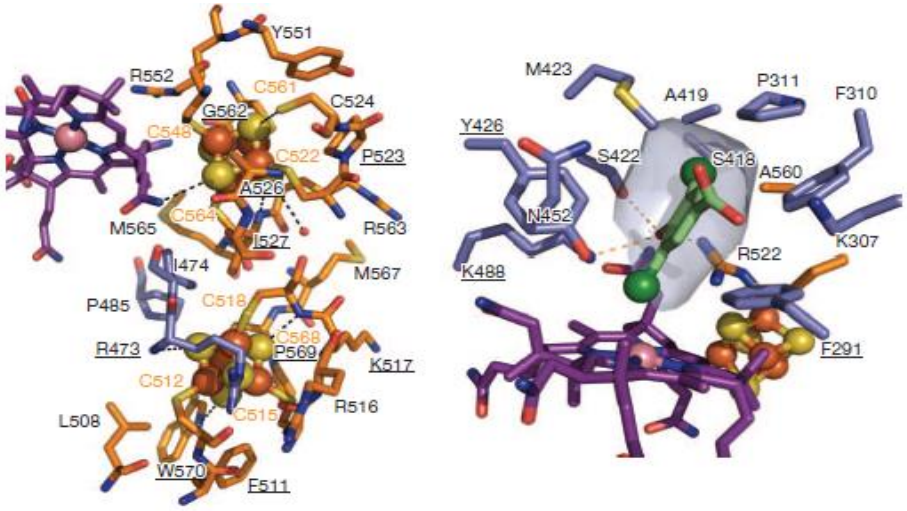


Figure 1.5 Structure of the NpRdhA protein binding to the corrin ring through 11 hydrogen bonds (figure on the left) (Payne *et al*., 2015). The figure on the right shows the docking of the substrate 3,5-dibromohydroxybenzoic acid at the active site of the enzyme complex. The region shows the active site cavity of the NpRdhA protein that binds to the substrate halogen atoms (Payne *et al*., 2015). Reprinted by permission from Macmillan Publishers Ltd: [Nature] (Payne KA, Quezada CP, Fisher K, Dunstan MS, Collins FA, Sjuts H, Levy C, Hay S, Rigby SEJ & Leys D (2015) Reductive dehalogenase structure suggests a mechanism for B12-dependent dehalogenation. *Nature* **7535**:513-516), copyright (2015). See appendix E2 for license terms from Nature Publishing Group.

Two potential catalytic mechanisms have been proposed for the enzyme (Payne et al., 2015). One possible catalytic mechanism for this structure is breakage of the halogen carbon bond homolytically after transfer of an electron from the [4F-4S] cluster or the Cob(I)alamin

forming the halogen-cob(II)alamin complex (Fig 1.6 Right). The conserved tyrosine 426 residue acts as the hydrogen donor to the substrate. A second possible mechanism for the dehalogenase catalytic reaction in this enzyme is the breakage of the halogen carbon bond heterolytically (Fig 1.6 Left).

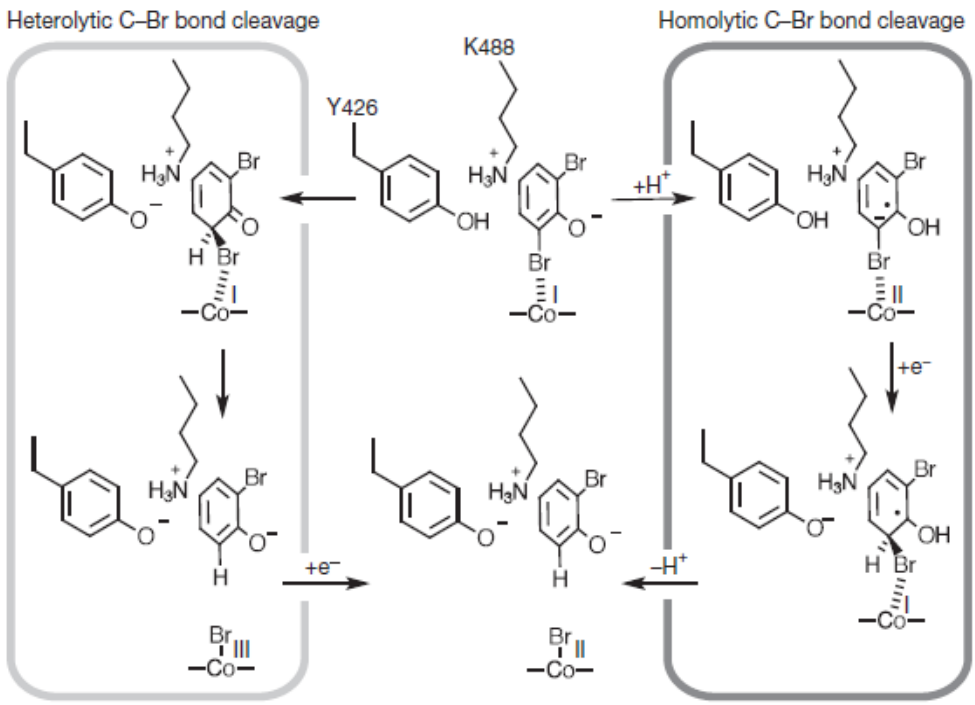


Figure 1.6 The two proposed dehalogenation mechanisms of catalysis by NpRdhA. The figure on the left shows the heterolytic breakage of the carbon-halogen bond and releasing the halogen-Cob(III)alamin complex. The figure on the right shows the homolytic breakage of the carbon-halogen bond after reuction by [4Fe-4S] cluster and releasing the halogen Cob (II) alamin complex. The conserved tyrosine residue at the active site functions as the hydrogen donor. Reprinted by permission from Macmillan Publishers Ltd: [Nature] (Payne KA, Quezada CP, Fisher K, Dunstan MS, Collins FA, Sjuts H, Levy C, Hay S, Rigby SEJ & Leys D (2015) Reductive dehalogenase structure suggests a mechanism for B12-dependent dehalogenation. *Nature* **7535**:513-516), copyright (2015). See appendix E2 for license terms from Nature Publishing Group.

1.8 Development of primers for amplification of reductive dehalogenases and study of dehalogenase gene expression.

Degenerate and specific primers were developed by Regeard *et al*. (2003) for the detection of chloroethene reductive dehalogenase. For the design of the degenerate primers, three chloroethene reductive dehalogenase and three chlorophenol reductive dehalogenase sequences that were available at EMBL under accession numbers AF228507 (tceA-De), AF022812 (pceA-Sm), AJ439607 (pceA-Dr), AF115542 (cprA-Dd),AY013365 (cprA-Dh), AF204275 (cprA-Dc) were aligned using several bioinformatic tools. The bioinformatic tools that were used for the alignment include T-Coffee (Notredame *et al*., 2000), Block Maker program (Henikoff et al., 1995), Motif algorithm (Smith *et al*., 1990) and the Gibbs Sampler Algorithm (Lawrence *et al*., 1993). With the help of these four Bioinformatics tools, seven conserved regions were identified and Consensus Degenerate Hybrid Ologonucleotide Primer design method (Rose *et al*., 1998) was used for the development of the degenerate primer sets. The degenerate primers were applied to the genomic DNA of *Sulfospirilium Multivorens* strain K, *Dehalobacter restrictus* strain PER-K23 and *Desulfitobacterium sp.* strain PCE1. Clones obtained using the degenerate primers were analyzed, and three novel putative *rdh* genes were detected in addition to the previously known chloroethene reductive dehalogenase genes. Specific primers for *pceA* gene was developed by using the Primer3 software (Rozen and Skaletsky 2000). The low copy number of the target genes was a drawback, and so a nested PCR method was developed in which amplification was carried out first by degenerate primers followed by specific primers to detect the low copy number genes.

The degenerate primers designed by Holliger *et al*. (2003) were further used by Marzorati *et al*. (2007) to amplify potential dehalogenating genes in a mixed culture that was

dechlorinating 1,2-dichloroethane to ethene. The PCR amplicons were further cloned and were used to develop specific primers for inverse and direct PCR. The inverse and the direct PCR approach using the specific primers designed produced a 7.4 Kb novel dehalogenating gene cluster. This gene cluster was found to have five open reading frames namely *dcaA*, *dcaB*, *dcaC*, *dcaT*, and *orf1*. The amino acid sequences of the dehalogenating cluster were aligned with other known dehalogenating enzyme sequences and it was found to be similar to the aligned sequences; it was concluded that the novel gene cluster encoded protein specific for 1,2-DCA dechlorination.

Krajmalnik-Brown *et al*. (2004) developed degenerate primers for amplification of a gene encoding a vinyl chloride reductase from mixed and pure cultures of *Dehaloccoides mccartyi* using the Clustal W software. The twin arginine motif (Tat signal sequence) and a conserved region internal to *Dehalococcoides rdhB* genes were identified and used for the development of degenerate primer sets RRF2 and BIR. This primer set when applied on the cDNA of mixed and pure cultures of *Dehaloccoides mccartyi* strain BAV1 gave an amplicon of the correct expected size of 1700 bp showing that both the *rdhA* and *rdhB* genes were transcribed together. A clone library was extablished using the PCR product and this lead to the identification 7 new putative reductive dehalogenases in strain BAV1 in pure and mixed cultures.

The degenerate primer set RRF2 and BIR was again used by Waller *et al*. (2005) to amplify and study the differential expression of reductive dehalogenase genes in *Dehaloccoides* containing mixed culture KB1. The degenerate primer sets were used for the amplification of both genomic DNA and also cDNA for the transcriptional analysis. These genes were also amplified in the transcriptional study, and these genes have shown remarkable similarities with the protein sequences of previously known reductive dehalogenases. In the transcriptional study

of the mixed KB1 cultures it was shown that multiple reductive dehalogenase were transcribed at the same time during the dechlorination of a single compound. Upregulation of as many as 29 *rdhA* genes were observed during the dechlorination of 1,2,3-trichlorobenzene and 1,2,4-trichlorobenzene in pure cultures of *Dehalococcoides mccartyi* strain CBDB1. Degenerate primers were designed for amplification of all 32 *rdhA* genes from the genome of *Dehalococcoides mccartyi* strain CBDB1. RT-PCR using these primers along with tRFLP showed the upregulation of multiple *rdhA* genes (Wagner *et al*., 2009). Evidence of upregulation of multiple genes encoding reductive dehalogenase proteins has also been observed in *Dehalococcoides mccartyi* strain 195 during the dechlorination of tetrachloroethene, trichloroethene, and 2,3-dichlorophenol (Fung *et al*., 2007). This study showed the highest transcript levels for the genes encoding PceA and TceA among all the transcripts deteced. The high translational levels of PceA and TceA proteins were also detected by proteomic analysis.

1.9 Regulatory genes associated with dehalogenase genes

The majority of the *rdhA* genes in *Dehalococcoides mccartyi* are associated with either two component system (TCS) regulators or MarR regulators (Kube *et al*., 2005; Seshadri *et al*., 2005). The two component systems (TCS) regulators consist of a histidine kinase and a response regulator. The histidine kinase contains transmembrane domains and is responsible for transfer of signals to the response regulator through a complex autophosphorylation pathway (Galperin, 2006). MarR proteins, on the other hand, often function as repressor proteins and regulate gene expression in response to change in environment, availability of antibiotics, and in response to aromatic compound metabolism (Wilkinson *et al*., 2006). MarR proteins have a helix-turn-helix DNA recognition fold, and in the presence of an aromatic ligand, they are typically released from the promoter sequences thereby inducing gene expression (Wilkinson *et al*., 2006).

Table 1.1 Number of *rdh* genes and associated regulated genes from *Dehalococcoides mccartyi* genomes (Wagner *et al*., 2013).

		Genes encoding TCS systems regulators		Genes encoding MarR regulators	
strain	*rdhAB* genes	total	Adjacent to *rdhA* genes	total	Adjacent to *rdhA* genes
CBDB1	32	24	14	16	10
195	17	19	10	8	3
BAV1	11	13	5	4	1
GT	20	3	1	12	7
VS	36	21	11	14	10

It is assumed that both TCS systems regulators and MarR regulators act as regulatory genes for the reductive dehalogenases. It has been proposed in *Dehalococcoides mccartyi* strain CBDB1 that the MarR protein (encoded by the gene locus *CbdbA* 1625) acts as a repressor. Promoter probe assays have demonstrated that the MarR protein plays a role in regulation of the *rdhA* genes (Wagner *et al*., 2013).

1.10 Previous studies of heterologous expression of reductive dehalogenases

The first attempts to heterologously express a functional reductive dehalogenase was reported by Neumann *et al*. (1998). The tetrachloroethene reductive dehalogenase *pceA* from *Dehalospirillum multivorans* was expressed nonfunctionally and in the insoluble fraction in *E. coli* and only in the presence of rare *E. coli* $tRNA_4^{ARG}$. The presence of the *E. coli* $tRNA_4^{ARG}$ compensated for the difference in codon usage in *D. multivorans* and *E. coli*. The reason for the production of a non-catalytically-active enzyme was attributed to the fact that the recombinant protein expressed in *E. coli* was an unprocessed PCE dehalogenase. The other reason presumed for the production of inactive protein was that the Fe/S clusters and the corrinoid cofactors were not incorporated in the protein since *E. coli* is not able to synthesize corrinoids.

The *pceA* gene from *Desulfitobacterium sp.* strain Y51 was also expressed in a non-catlytically-active form in *E. coli* (Suyama *et al*., 2002). Isopropyl β-D-1-thiogalactopyranoside induced cell extracts of recombinant *E. coli* cells harbouring the *pceA* in a pET expression system gene failed to show dehalogenase activity under both aerobic and anaerobic conditions. Attempts to refold the protein from the insoluble fraction did not result in the production of a functionally active protein, and again it was assumed that the absence of a *de novo* cobamide biosynthesis pathway in *E. coli* led to the production of a non-functional enzyme.

The first instance of heterologous production of a soluble reductive dehalogenase was the *pceA* gene from *Dehalobacter restrictus* (Sjuts *et al*., 2012). The soluble protein was expressed in an *E. coli* host system by fusing the PceA with a strep tag as a solubility tag (Trigger factor) and by expressing the protein at cold temperatures. The fusion protein was purified using a StrepTactin Sepharose resin and the 4Fe-4S clusters were incorporated into the protein by incubation with NaS and $FeCl_3$ under anaerobic conditions. The protein, however, was still enzymatically inactive, and again it was assumed that the right cobamide was not available for the enzyme to be functional or the 4Fe-4S was not completely reconstituted to produce the functional enzyme.

The first report in the literature of successful heterologous expression of a catalytically active reductive dehalogenase came out very recently in (Mac Nelly *et al*., 2014), and for the first time, a genetic system for the production of a functional reductive dehalogenase was identified. *Shimwellia blattae* (ATCC 33430) was used as the host for the expression of functional reductive dehalogenase because of its ability to synthesize cobamides *de novo* and hence was preferred in place of the standard host *E. coli*. The tetrachloethene reductive dehalogenase (PceA) of the *Desulfitobacterium hafniense* strain Y51 was functionally expressed

in the *S. blattae* cells. *S. blattae* cells were grown under anoxic conditions for the production of the recombinant protein, and glycerol was used as the sole energy source. It was observed that the production of the soluble enzyme increased when the trigger factor chaperone PceT located downstream was co expressed along with the *pceA* gene. After the PceA reductive dehalogenase was functionally expressed, additional experiments were performed using the *rdhA3* of *D. hafniense* DCB-2, and catalytically active RdhA3 was again expressed. The catalytic activity of the PceA enzyme was notably increased when 5,6-dimethylbenzimidazole (DMB) and hydroxocobalamine were added to the growth medium of the recombinant enzyme-producing cells, suggesting their roles as cofactors for the processing of the enzyme. Also, it was reported that active form of the enzyme was also expressed in the absence of the trigger factor (PceT); hence, this system could possibly be used for the functional expression of other dehalogenase genes with unknown substrate specificity existing in the present databases.

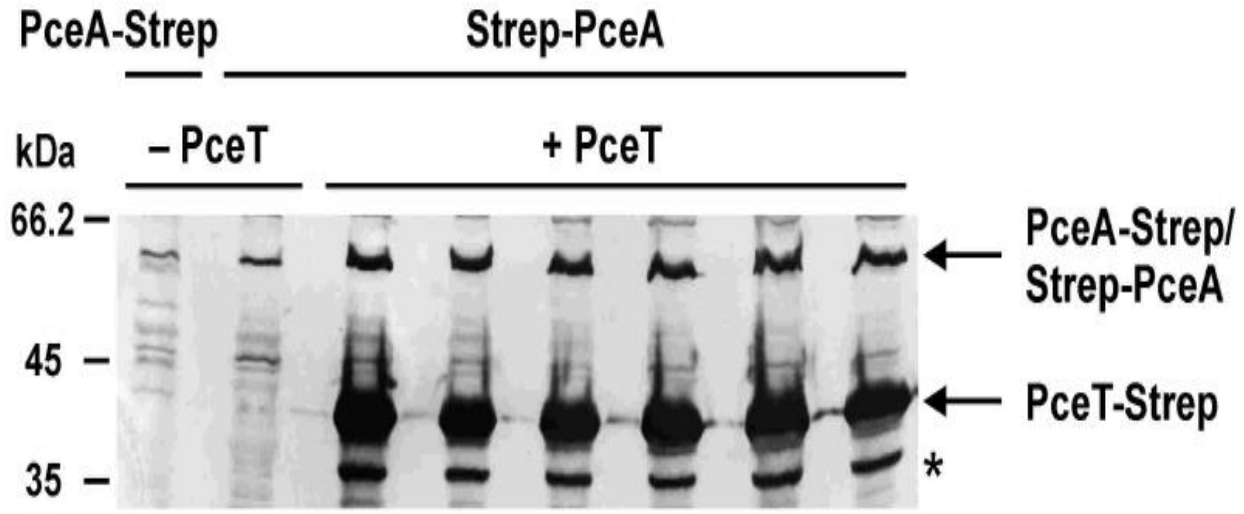


Figure 1.7 Heterologous production of soluble functional PceA from *S. blattae* (McNelly *et al*., 2014, reprinted with permission from ASM). Strep tagged soluble PceA detected by immunoblot. Solubility of the PceA protein was found to be more when the PceT protein was co expressed.

1.11 Research Objectives

The first objective of research described in this dissertation was to gain insights into the extent of reductive dehalogenase gene expression in *Dehalogenimonas lykanthroporepellens* BL-DC-9^{T} during dechlorination of 1,2-dichloroethane (1,2-DCA), 1,2-dichloropropane (1,2-DCP), and 1,2,3-trichloropropane (1,2,3-TCP). The second research objective was to develop a method for functional heterologous expression of an active dehalogenase enzyme using a suitable host.

Chapter 2. Expression of putative reductive dehalogenase genes from BL-DC-9^T during dechlorination Of 1,2-DCA, 1,2-DCP and 1,2,3-TCP[1]

2.1 Introduction

As described in Chapter 1, the genome of *D. lykanthroporepellens* strain BL-DC-9^T (GenBank accession no. CP002084) contains 25 loci having reductive dehalogenase homologous (*rdh*) genes scattered throughout the chromosome (Siddaramappa *et al*., 2012) (Table 2.1). For genera other than *Dehalogenimonas* (for which only one functional enzyme has yet been characterized) (Padilla-Crespo *et al.*, 2014), the genes encoding enzymes characterized to date that are involved in catalyzing the anaerobic reductive dehalogenation of chlorinated solvents are organized in *rdhAB* operons encoding two components: a 50-65 kDa protein (RdhA) that functions as a reductive dehalogenase and a ~10 kDa hydrophobic protein with transmembrane helices (RdhB*)* that is thought to anchor the RdhA to the cytoplasmic membrane (Magnuson *et al*., 2000; Maillard *et al*., 2003; Smidt *et al*., 2004; Futagami *et al*., 2008). Genera sharing this feature include *Sulfurospirillum*, *Dehalobacter*, *Desulfitobacterium*, and *Dehalococcoides* (Miller *et al*., 1998; Neumann *et al*., 1998; Magnuson *et al*., 2000; Maillard *et al*., 2003; Suyama *et al*., 2002; Müller *et al*., 2004; Krajmalnik-Brown *et al*., 2004; He *et al*., 2005; Padilla-Crespo *et al*., 2014). Previously characterized RdhA proteins also contain a twin-arginine motif near the N-terminus that is thought to be associated with the twin-arginine translocation (TAT) system, a specialized system involved in the secretion of folded proteins across the bacterial inner membrane into the periplasmic space (Stanley *et al*., 2000). Two conserved motifs each

[1] Portions of this chapter previously appeared as [Kalpataru Mukherjee1 , Kimberly S. Bowman1,2, Fred A. Rainey1,3, Shivakumara Siddaramappa4,5, Jean F. Challacombe4 & William M. Moe2 Dehalogenimonas lykanthroporepellens BL-DC-9T simultaneously transcribes many rdhA genes during organohalide respiration with 1,2-DCA, 1,2-DCP, and 1,2,3-TCP as electron acceptors FEMS Microbiol Lett. 2014 May;354(2):111-8.

containing four cysteine residues, a feature associated with binding iron-sulfur clusters (Bruschi & Guerlesquin, 2000) have also been reported to occur near the C-terminus of previously characterized rdhA proteins (Magnuson *et al*., 2000, Maillard *et al*., 2003; Hug *et al*., 2013).

All of the *rdhB* code for proteins predicted to contain two or three transmembrane helices (Table 2.2), consistent with observations for putative dehalogenase membrane-anchoring proteins in *Dehalococcoides mccartyi* strains (Kube *et al*., 2005; Seshadri *et al*., 2005; McMurdie *et al*., 2009; Padilla-Crespo *et al*., 2014).

The majority of *rdh* genes in *Dehalococcoides* strains are associated with MarR or two component system transcriptional regulators. For example, *Dehalococcoides mccartyi* strain 195 has transcriptional regulators for 17 of the *rdh* genes present in its genome while *Dehalococcoides* strain CBDB1 has transcriptional regulators for all 32 of its *rdh genes* (Kube *et al*., 2005; Seshadri *et al*., 2005). In contrast, in the genome sequence of *Dehalogenimonas lykanthroporepellens* BL-DC-9^T, such transcriptional regulators are associated with only four of the 25 *rdhA* genes. These include one MarR type transcriptional regulator and three two-component system transcriptional regulators (Table 2.3).

The broad goal of experiments presented in this chapter was to determine which of the *rdh* genes are expressed by *D. lykanthroporepellens* strain BL-DC-9^T during the anaerobic reductive dechlorination of different compounds and investigate genes that may regulate their expression. Oligonucleotide primers were designed and PCR protocols that allow PCR amplification of all 25 *rdhA* genes (Table 2.4) and transcription factors (Table 2.5) were developed. Subsequently, the PCR primers and protocols were employed in conjunction with reverse transcriptase PCR (RT-PCR) to determine which genes were expressed during growth of BL-DC-9^T with the chlorinated solvents 1,2-DCA, 1,2-DCP, and 1,2,3-TCP.

Table 2.1: Characteristics of putative reductive dehalogenases (rdhA) of *D. lykanthroporepellens* BL-DC-9^T

Locus tag, protein	Chromosomal location		ORF Size (bp)	GC %	IS element association	Tat signal sequence	Fe-S binding motif #1	Fe-S binding motif #2	Closest homolog Locus, protein, identity[a]	Cognate *rdhB*
	Start (bp)	Stop (bp)								
Dehly_0068, 460 aa	65748	64366	1383	57.27	None	+	$CX_2CX_2CX_3C$	$CX_9CX_4CX_3C$	dcmb_1434, 482aa, 40%	None
Dehly_0069, 454 aa[b]	67393	66029	1365	60.66	Dehly_0070	-	$CX_2CX_2CX_3C$	$CX_9CX_4CX_3C$	cbdbA1491, 482 aa, 36%	None
Dehly_0075, 300 aa[b]	75099	74197	903	55.92	Dehly_0076	-	$CX_2CX_2CX_3C$	$CX_9CX_4CX_3C$	cbdbA1542, 487 aa,39%	None
Dehly_0121, 469 aa	115115	116524	1410	51.21	None	+	$CX_2CX_2CX_3C$	$CX_8CX_2CX_3C$	DehaBAV1_0281, 470 aa, 39%	None
Dehly_0156, 467 aa	153539	154942	1404	59.19	None	+	$CX_2CX_2CX_3C$	$CX_2CX_2CX_3C$	DhcVS_1340, 480aa,43%	Dehly_0157
Dehly_0274, 472 aa	262130	263548	1419	45.38	None	+	$CX_2CX_2CX_3C$	$CX_8CX_2CX_3C$	DehaBAV1_0281, 470aa, 45%	None
Dehly_0275, 493 aa	263669	265150	1482	45.34	None	+	$CX_2CX_2CX_3C$	$CX_2CX_2CX_3C$	DehaBAV1_0284, 496 aa, 78%	Dehly_0276
Dehly_0283, 476 aa	269792	268362	1431	43.89	Dehly_0284	+	$CX_2CX_2CX_3C$	$CX_{14}CX_2CX_3C$	cbdbA84, 488 aa, 38%	None
Dehly_0479, 339 aa[b]	447957	448976	1020	49.71	Dehly_0480	+	None	None	cbdbA1563, 465 aa,40%	None
Dehly_0849, 475 aa	830820	832247	1428	54.62	None	+	$CX_2CX_2CX_3C$	$CX_9CX_4CX_3C$	cbdbA1542, 487 aa,38%	None
Dehly_0910, 464 aa	896375	894981	1395	53.62	None	+	$CX_2CX_2CX_3C$	$CX_8CX_2CX_3C$	DhcVS_1342, 473 aa, 59%	None
Dehly_1054, 473 aa	1050007	1048586	1422	54.50	None	+	$CX_2CX_2CX_3C$	$CX_2CX_2CX_3C$	DhcVS_1340, 480 aa,44%	Dehly_1053
Dehly_1148, 462 aa	1129672	1128284	1389	51.33	None	+	$CX_2CX_2CX_3C$	$CX_{11}CX_2CX_3C$	GY50_1378, 508aa, 36%	None
Dehly_1152, 458 aa	1133813	1132437	1377	56.21	None	+	$CX_2CX_2CX_3C$	$CX_2CX_2CX_3C$	DehaBAV1_0284, 496 aa, 45%	Dehly_1151
Dehly_1328, 476 aa	1305989	1307419	1431	47.87	None	+	$CX_2CX_2CX_3C$	$CX_{10}CX_2CX_3C$	cbdbA1563, 465aa, 43%	None
Dehly_1355, 474 aa	1329395	1330819	1425	47.09	Dehly_1354	+	$CX_2CX_2CX_3C$	$CX_{11}CX_2CX_3C$	cbdbA1542, 487 aa, 41%	None
Dehly_1514, 455 aa	1485472	1484105	1368	44.37	Dehly_1515	+	$CX_2CX_2CX_3C$	$CX_{16}CX_2CX_3C$	cbdbA1563, 465 aa, 35%	None
Dehly_1520, 129 aa	1493625	1494014	390	37.44	Dehly_1521	-	None	None	DehaBAV1_0281, 470 aa, 32%	None
Dehly_1523, 340 aa	1495262	1496284	1023	43.01	Dehly_1522	-	$CX_2CX_2CX_3C$	$CX_{19}CX_2CX_3C$	DhcVS_1353, 514 aa, 31%	None
Dehly_1524, 482 aa	1496492	1497940	1449	44.93	Dehly_1526	+	$CX_2CX_2CX_3C$	$CX_{21}CX_2CX_3C$	AGS15112, 484 aa, 92%	Dehly_1525
Dehly_1530, 473 aa	1503758	1502337	1422	49.23	Dehly_1529	+	$CX_2CX_2CX_3C$	$CX_{11}CX_7CX_3C$	dcmb_1434, 482 aa, 39%	None
Dehly_1534, 295 aa[b]	1507895	1507008	888	54.17	Dehly_1535	-	$CX_2CX_2CX_3C$	$CX_2CX_2CX_3C$	DehaBAV1_0284, 496 aa, 49%	Dehly_1533
Dehly_1540, 465 aa	1514181	1515578	1398	47.57	Dehly_1542	+	$CX_2CX_2CX_3C$	$CX_2CX_2CX_3C$	DehaBAV1_0281, 470 aa, 35%	None
Dehly_1541, 51aa[c]	1515618	1515773	156	59.62	None	-	None	None	DhcVS_1421, 475 aa, 55%	None
Dehly_1582, 452 aa	1547392	1546034	1359	60.56	Dehly_1583	-	$CX_2CX_2CX_3C$	$CX_8CX_2CX_3C$	DehaBAV1_0281, 470 aa, 42%	None

[a]Identity across the entire length of the query protein

[b]Annotated as pseudogenes

[c]Truncated ORF ,annotated as hypothetical proteins

Table 2.2. Characteristics of putative reductive dehalogenase membrane anchoring proteins (rdhB) of *D. lykanthroporepellens* BL-DC-9^T.

Locus Tag, Protein	Chromosomal Location		ORF size (bp)	GC %	TM[a]	Cognate *rdhA*	Closest homolog Locus tag, protein, identity
	Start	Stop					
Dehly_0157, 93 aa	154978	155259	282	61.35	3	Dehly_0156	GY50_1404, 95 aa, 43%
Dehly_0276, 91 aa	265171	265446	276	40.58	3	Dehly_0275	DehaBAV1_0283, 88 aa,72%
Dehly_1053, 94 aa	1048546	1048262	285	57.19	3	Dehly_1054	DhcVS_1363, 95 aa, 41%
Dehly_1151, 90 aa	1132415	1132143	273	51.28	3	Dehly_1152	DhcVS_1307, 89 aa, 43%
Dehly_1504, 88 aa	1476013	1476279	267	57.30	3	None	dcmb_118, 90 aa, 63%
Dehly_1525, 72 aa	1497959	1498177	219	35.16	2	Dehly_1524	DET0163, 76 aa, 47%
Dehly_1533, 88 aa	1506976	1506710	267	48.69	3	Dehly_1534	btf_1459, 89 aa,51%

[a] Number of transmembrane helices as predicted by TMHMM2.0 (Krogh *et al*., 2001)

Table 2.3: Characteristics of putative transcriptional regulators associated with *rdhA* genes.

Locus Tag Protein	Type of regulatory element	Associated *rdhA*
Dehly_0120	MarR	Dehly_0121
Dehly_0270	Two component system transcriptional regulator	Dehly_0274
Dehly_1149	Two component system transcriptional regulator	Dehly_1148, Dehly_1152
Dehly_1531	Two component system transcriptional regulator	Dehly_1530

2.2 Materials and Methods

2.2.1 Primer design and PCR protocols

Oligonucleotide primers specific for the *rdhA* and *rdhB* genes of *D. lykanthroporepellens* BL-DC-9^T were designed using a combination of the Primer3 software package (Rozen & Skaletsky, 2000)/ Blast at NCBI/Primer-Blast and also by manual selection after alignment of all *rdh* genes in the genome using the BioEdit 7.05.3 software package (Hall, 1999). The Primer 3 software enabled the screening of hairpin loops and primer dimers during primer design. Also

the primer specificity was checked using Blast. For the six *rdhA* genes with cognate *rdhB* genes, primers were selected to span the two genes (i.e., forward primer targeting location in *rdhA* and reverse primer binding site in *rdhB*). Primers were likewise designed for the amplification of all four regulatory genes.

Table 2.4. Primers and annealing temperatures used for the detection of *rdh* gene expression.

Target locus tag	Target Gene ID	Primer ID	Primer sequence(5'-3')	Expected product size (bp)	Annealing temp (°C)
Dehly_0068	9389745	Dehly_0068F	TCATGAAAGTCCTCGGCCTG	434	55
		Dehly_0068R	CTCTTCCTTGGTACCTTCCC		
Dehly_0069	9389746	Dehly_0069R	TAGCCGAGGGTGCGGACGAA	360	60
		Dehly_0069F	TATCCCCCGCTGGAACGGCA		
Dehly_0075	75099	Dehly_0075R	GCGTAGTAGCCGAGGGAAGT	270	58
		Dehly_0075F	CCATGAGGACACCACCGTGAAG		
Dehly_0121	9389797	Dehly_0121F	GACCGGCGCAGACAGTTGCT	462	60
		Dehly_0121R	GACGGCATCATACCGCCCGG		
Dehly_0156-Dehly_0157	9389835, 9389836	Dehly_0156F	ACTGTGCCGGTTGTCAGGCG	503	65
		Dehly_0157R	AGCCGAAGAGGCCCAGCAGA		
Dehly_0274	9389954	Dehly_0274F	GGCCGAACGCTCCATGGGAT	345	63
		Dehly_0274R	CCATGCCGAAACCACCCGGA		
Dehly_0275-Dehly_0276	9389955, 9389956	Dehly_0275F	TCCACATTGCCCGGTTTGCCA	419	60
		Dehly_0275R	AGTCCCGCCGTTGGCTCATA		
Dehly_0283	9389963	Dehly_0283R	GTCGCACTCCACCTTGGTAA	438	60
		Dehly_0283F	ACTTGGATTGTCTGGGGCAG		
Dehly_0479	447957	Dehly_0479F	TCGCCCAAGGGGGACTGGAA	410	55
		Dehly_0479R	ACGCCTACGTCCCTGGCTCC		
Dehly_0849	9390556	Dehly_0849F	ATACCAGGCGGTCGGCCAGT	372	60
		Dehly_0849R	CCTGAGGTGCAGGCGTTGGG		
Dehly_0910	9390618	Dehly_0910R	CGGCATCAAGCACTGGCCCA	244	63
		Dehly_0910F	ATCCCGGTTCCACCAGGCCA		
Dehly_1053-Dehly_1054	9390762, 9390763	Dehly_1053F	ACGACCCCACCTGGGACACC	608	67
		Dehly_1053R	CCAACCGGAAACGGCCAGCA		
Dehly_1148	9390857	Dehly_1148F	GGCACTTTGGCGGGTCGGAA	551	60
		Dehly_1148R	GGCGTCCTCGGTGCTGATCG		
Dehly_1151-Dehly_1152	9390860, 9390861	Dehly_1152F	TTGCCCGCATTGTCCGCAGT	349	65
		Dehly_1152R	GGCGCCCATCAGCCATTCGT		
Dehly_1328	9391042	Dehly_1328F	CCAACCACCGAAGTTGACTGG	347	60
		Dehly_1328R	GCCCCATAAAAGCGTGCGGC		
Dehly_1355	9391069	Dehly_1355F	GGGATGGCAGTCAACCCGGC	464	60
		Dehly_1355R	AACGCCGCCCGAATCATCCG		
Dehly_1514	9391237	Dehly_1514F	AGCCGGACTAGGTGCAGCCT	310	58
		Dehly_1514R	TGCGGTACGGCCGCCAAATT		
Dehly_1520	9391243	Dehly_1520F	TGGACCAGATGTTGAGTGCATCGG	217	68
		Dehly_1520R	TGCAGAATGCCATTTGTTTCGCGG		
Dehly_1523	9391246	Dehly_1523F	TGCCACGCTGGGAAGGCAAG	892	68
		Dehly_152 R	TCCGGCAGGTCGATGTCCCA		
Dehly_1524-Dehly_1525	9391247, 9391248	Dehly_1524F	TGGCACCAGTACCTGCCCGCTAT	168	58
		Dehly_1524R	GACAACGCCAGGCTGAGATTCTGC		
Dehly_1530	9391253	Dehly_1530F	AGTTGCCCAAGCGGCAAGGG	196	60
		Dehly_1530R	ATCCACCGCTTCTGCGCCAC		

(Table 2.4 Continued)

Target locus tag	Target Gene ID	Primer ID	Primer sequence(5'-3')	Expected product size (bp)	Annealing temp (°C)
Dehly_1533-Dehly_1534	9391256, 9391257	Dehly_1533F	ATGGGGTGTTACCCCGCGT	677	65
		Dehly_1533R	GCCGATGGTGGCGGCTGTTT		
Dehly_1540	9391263	Dehly_1540F	TTGTGGGCGCTGGTAACGCT	387	50
		Dehly_1540R	CCACGCGGACGATAGTGCCG		
Dehly_1541	9391264	Dehly_1541F	TTCATGAAAGGGCTTGGGCTTGCC	115	50
		Dehly_1541R	TCACACGGTAGTTCGTTTCCGG		
Dehly_1582	9391305	Dehly_1582F	GGATGACAGCTATCCCGGTG	437	55
		Dehly_1582R	AGTTGACGTACCAGCCCTTG		

Primers were tested using genomic DNA of *D. lykanthroporepellens* BL-DC-9^T (=JCM 15061^T = ATCC BAA-1523^T), extracted using an UltraClean Microbial DNA isolation kit (MoBio Laboratories, CA), as template in PCR reactions. PCR was carried out in 25 µL reaction volumes using 10 ng of genomic DNA (as quantified by NanoDrop 2000 spectrophotometer, Thermo Fisher Scientific), 0.4 mM dNTPs, 1 U Platinum *Taq* polymerase (Invitrogen) and 0.4 µM of forward and reverse primers. Amplifications were carried out with initial denaturation at 95°C for 2 min, thirty cycles consisting of 95°C for 15 sec, primer specific annealing temperature for 30 sec (Tables 2.4 and 2.5), and extension at 68°C for 60 sec. A final extension was carried out at 68°C for five min.

Table 2.5: Primers and annealing temperature used for study of transcription factor gene expression.

Target locus tag	Primer ID	Primer sequence (5'-3')	Expected product size (bp)	Annealing temp (°C)
Dehly_0120	Dehly_0120F	GTCATGCGGATGTTTCCGTG	225bp	59
	Dehly_0120R	AAGGGGTTTGTCAGCTCTCG		
Dehly_0270	Dehly_0270F	AGCAGTTTCGCCTTCTCCAA	143bp	59
	Dehly_0270R	CGACTACCAGGATGTCAGGC		
Dehly_1149	Dehly_1149F	CCGAGTCCGGAGACGTTATG	622bp	59
	Dehly_1149R	ACGATCCCGGAGTAGTGGAA		
Dehly_1531	Dehly_1531F	AATCCAGCACCTGCTTCTCC	185bp	59
	Dehly_1531R	CAAAAACACACCGCGCAAAC		

Following PCR, amplification products were separated via electrophoresis in a 2% (m/v) agarose gel and visualized using ethidium bromide staining. A 100 bp ladder (Sigma-Aldrich) was used to estimate the sizes of PCR amplicons. Production of amplicons of the size expected based on the genome sequence was taken as an initial indication of success. PCR products were subsequently sequenced to verify that the primers and PCR protocols allowed unique detection of the intended gene targets. Sequencing was performed using the BigDye Terminator version 3.1 (Applied Biosystems) in conjunction with an ABI prism 3130 sequencer.

To evaluate the detection limits for the PCR primers and thermal protocols (Tables 2.4 and 2.5), genomic DNA from *D. lykanthroporepellens* BL-DC-9^T quantified using a Qubit dsDNA HS assay kit (Invitrogen) in conjunction with a Qubit2 fluorometer was serially diluted (10 ng/µL to 0.00015 ng/µL) and then used as PCR template (with 1 µL template DNA per reaction). This corresponds to 5.8×10^6 to 8.7×10^1 genome copies/PCR. PCR conditions and amplicon detection were as described above.

2.2.2 Cell culture

Cultures for transcription experiments were grown in 25 mL glass serum bottles (Wheaton) sealed with butyl rubber stoppers and aluminum crimp caps. Each serum bottle contained 15 mL of anaerobic basal liquid medium with titanium citrate as reducing agent (Maness *et al*., 2012) and 10 mL gas headspace comprised of 10% CO_2, 80% N_2, and 10% H_2 (mol%). Triplicate bottles spiked with filter sterilized 1,2-DCP, 1,2-DCA, or 1,2,3-TCP to reach an aqueous phase concentration of 2 mM following dissolution and equilibration were inoculated (6% v/v) with *D. lykanthroporepellens* BL-DC-9^T. Incubation was in the dark at 30°C without mixing. Triplicate bottles prepared, inoculated, and incubated in identical fashion but without addition of chlorinated solvents were prepared at the same time as controls. Cultures used for

inoculation of bottles used in transcription experiments were grown on the same solvent prior to use as inoculum (e.g., the *D. lykanthroporepellens* BL-DC-9^T culture inoculated into bottles containing 2 mM 1,2-DCA was grown on 2 mM 1,2-DCA) and were sparged with a gas mixture comprised of 10% CO_2, 80% N_2, and 10% H_2 (mol%) to remove residual chlorinated compounds prior to use as an inoculum. Dechlorination was monitored by gas chromatography as described previously (Yan *et al*., 2009). Bottles were incubated until 25-65% of the parent chlorinated solvent was dechlorinated (see Table 3) prior to harvesting biomass.

2.2.3 RNA extraction and purification

Cells for RNA extraction were collected by centrifugation. Serum bottles were opened inside an anaerobic chamber (Coy) supplied with gas headspace comprised of 10% CO_2, 80% N_2, and 10% H_2 (mol %), and 15 mL culture volumes were transferred into plug sealed centrifugation tubes (Corning). The centrifuge tubes were removed from the anaerobic chamber and centrifuged at 5900×*g* for 30 min at 4°C. After briefly opening tubes to decant supernatant (< 1 min), the remaining cell pellets were immediately frozen by submerging tubes in liquid N_2. Frozen cells were resuspended in 1 mL of TRI reagent (Molecular Research Center), and RNA was extracted using the modified guanidium thiocyanate-phenol-chloroform RNA extraction method of Chomczynski *et al*. (2006). Extraction was with 200 µL chloroform and precipitation was with 500 µL isopropanol. For purification, the RNA pellet was washed with two rounds of 70% ethanol, and RNA was eluted in 30 µL of nuclease free water. The extracted RNA was subjected to two rounds of DNase treatment using an RQ1 RNase-free DNase kit (Promega) before aliquots of RNA extract from each sample were analyzed using an RNA 6000 Pico kit (Agilent) in conjunction with an Agilent 2100 Bioanalyzer to assess RNA quality, quantify concentrations, and calculate the RNA integrity number (RIN, Schroeder *et al*., 2006).

2.2.4 Reverse transcription experiments

End point RT-PCR was carried out using the *rdh*-specific primers (Table 2.4) and primers targeting the seven regulatory genes (Table 2.5) the same day that RNA was extracted. In the RT-PCR reaction, cDNA synthesis and PCR were carried out in the same tube using the SuperScript III one step RT-PCR system with Platinum *Taq* DNA polymerase kit (Invitrogen). 0.5 µL of DNase-treated RNA extract was used as template for each reaction. The RT-PCR thermal program was 55°C for 30 min for reverse transcription followed by initial denaturation at 95°C for 2 min and then thirty cycles of 95°C for 15 sec, primer specific annealing temperature for 30 sec (see Table 2.4 and 2.5) and extension at 68°C for 60 sec. A final extension was carried out at 68°C for five min. To check for the presence of genomic DNA contamination and as a negative control, in parallel, PCR reactions were carried out with 0.5 µL RNA extract serving as template in PCR using Platinum *Taq* DNA polymerase in the absence of reverse transcriptase enzyme.

The end point RT-PCR products and samples from control PCR reactions were analyzed by agarose gel electrophoresis. The detection of amplicons of the correct size based on the genome sequence in reverse transcribed but not direct PCR (i.e., lacking reverse transcription) was used as an indicator of gene expression. To verify amplicon identity, RT-PCR amplicons were excised from the electrophoresis gels, eluted using a MinElute gel extraction kit (Qiagen), and sequenced. The RT-PCR reactions were carried out in triplicate for each of the primer sets using cells harvested from triplicate bottles.

2.3 Results

2.3.1 Primer design and detection limits

Using genomic DNA as template, PCR products of the correct size, as predicted from the genome sequence, were observed at template DNA concentrations of 0.001 ng/µL and higher (Figures 2.1 and 2.2) but not at 0.00075 ng/µL and lower for all of the final PCR primer combinations and thermal protocols tested (Tables 2.4 and 2.5). This corresponds to a detection limit of 5.8×10^2 genome copies per PCR reaction. Sequencing of expected sized PCR products showed that the amplicons were exact copies of target loci present in the chromosome of *D. lykanthroporepellens* BL-DC-9^T.

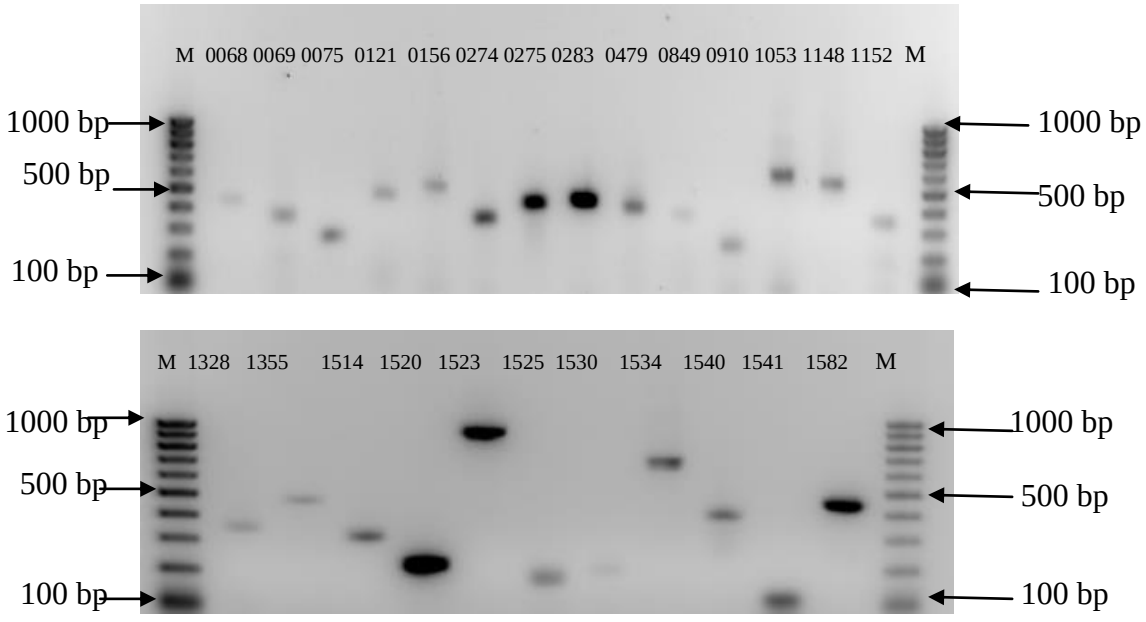


Figure 2.1. Agarose gel image showing detection of expected sized amplicons amplified using the designed *rdh* primers using DNA template concentration of 0.001 ng/µL. The molecular marker designated as M is DirectLoad 100 bp Low Ladder (Sigma-Aldrich).

2.3.2 Expression of *rdhA* genes during dechlorination of different substrates

Reductive dechlorination was observed in the bottles inoculated with strain BL-DC-9^T and spiked with all three of the chlorinated compounds provided as electron acceptors following pathways described previously (i.e., 1,2-DCA transformed to ethene, 1,2-DCP transformed to propene, and 1,2,3-TCP transformed to allyl chloride, Moe *et al*., 2009; Yan *et al*., 2009). The amount of dechlorination observed at the time when bottles were sacrificed is provided in Table 2.6. The extracted RNA was of high quality, with RNA integrity numbers (RIN) ranging from 7.5-9.6 (Table 2.6). Appendix A shows the results from the Bioanalyzer assay to check the RNA quality.

For BL-DC-9^T cells harvested from all replicates of cultures that were actively dechlorinating each of the three chlorinated alkanes (1, 2-DCA, 1,2-DCP and 1,2,3-TCP), amplicons were detected in RT-PCR with 19 of the 25 *rdh*-specific primer sets (Table 2.6). No amplicons were detected following RT-PCR with RNA extracted from triplicate bottles prepared at the same time and incubated under identical conditions but without the presence of a halogenated electron acceptor. No PCR amplicons were observed in any of the tests when RNA extract was used as template in PCR without reverse transcription (conducted to detect the presence of genomic DNA contamination in the RNA extracts). Agarose gel images supporting these results are shown in Appendix B. The amplicons detected following the RT-PCR were sequenced using the BigDye Terminator version 3.1 (Applied Biosystems) in conjunction with an ABI prism 3130 sequencer. The sequences were aligned with the target genes from the genome sequence. Sequence alignments using the sequenced RT-PCR result showed that the right targets were detected in the RT-PCR experiments targetting the *rdh* genes. The RT-PCR

result for the target Dehly_1541 had two bands. Both the bands were excized from the gel and sequenced and the expected sized band was identified as the correct target after sequencing.

Of the 17 "full length" *rdhA* genes, 13 were transcribed during dechlorination of all three halogenated solvents (Dehly_0121, Dehly_0156/Dehly_0157, Dehly_0274, Dehly_0275/Dehly_0276, Dehly_0849, Dehly_0910, Dehly_1053/Dehly_1054, Dehly_1148, Dehly_1152, Dehly_1328, Dehly_1514, Dehly_1524/Dehly_1525, and Dehly_1530). Transcripts from the remaining four full-length *rdhA* genes (Dehly_0068, Dehly_0283, Dehly_1355, Dehly_1540) were not detected.

Of the truncated *rdhA* genes, transcripts were detected for those lacking a predicted twin-arginine sequence at the N-terminus (Dehly_0069, Dehly_1523, Dehly_1582), those lacking predicted iron-sulfur cluster binding motifs at the C-terminus (Dehly_0479), and one that lacked both a predicted twin-arginine sequence at the n-terminus and iron-sulfur cluster binding motifs at the C-terminus (Dehly_1520).

Of the six loci where *rdhA* genes are accompanied by congnate *rdhB* genes, amplicons were detected using five primer sets that targeted regions spanning the *rdhAB* genes (Dehly_0156/Dehly_0157, Dehly_0275/Dehly_0276, Dehly_1053/Dehly_1054, Dehly_1151/Dehly_1152, Dehly_1524/Dehly_1525), indicating that the *rdhAB* genes are co-transcribed by BL-DC-9^{T}. The detected transcripts included the low mol% G+C *rdhB* gene predicted to code for a putative membrane anchoring protein containing two as opposed to three transmembrate helices (Dehly_1524). Amplification products were not detected following RT-PCR with primers targeting the other *rdhAB* gene combination (Dehly_1533/Dehly_1534), indicating that this locus was not highly expressed or that the genes are not co-transcribed.

2.3.3 Expression of the transcription factors during dechlorination of different substrates

All four transcription factors (Dehly_0120, Dehly_0270, Dehly_1149, and Dehly_1531) were detected in the reverse transcriptase experiments from BL-DC-9^T cells that were actively dechlorinating 1,2-DCA, 1,2-DCP, and 1,2,3-TCP. No amplicons were detected following RT-PCR with RNA extracted from triplicate bottles prepared at the same time and incubated under identical conditions but without the presence of halogenated electron acceptors. No PCR amplicons were detected in any of the tests when RNA extract was used as template in PCR without reverse transcription. The agarose gel images supporting these results are presented in appendix C. The amplicons detected following the RT-PCR targetting the regulatory genes were sequenced using the BigDye Terminator version 3.1 (Applied Biosystems) in conjunction with an ABI prism 3130 sequencer. The sequences were aligned with the target genes from the genome sequence. Sequence alignments using the sequenced RT-PCR result showed that the right targets were detected in the RT-PCR experiments targetting the regulatory genes.

2.4 Discussion

It is clear from the RT-PCR results presented here that *D. lykanthroporepellens* BL-DC-9^T simultaneously transcribes many *rdh* genes during the reductive dehalogenation of all three chlorinated alkanes tested. Simultaneous transcription of multiple *rdhA* genes in response to a single chlorinated compound has been reported previously for strains of the related genus *Dehalococcoides* (Waller *et al*., 2005; Fung *et al*., 2007; Rowe *et al*., 2008; Johnson *et al*., 2008; Rahm & Richardson, 2008; Wagner *et al*., 2009; Wagner *et al*., 2013).

Table 2.6: Summary of results from RT-PCR experiments for *rdh* genes with BL-DC-9^T grown with 1,2-DCA, 1,2-DCP, and 1,2,3-TCP.

Yes=PCR amplicon of appropriate size was observed in electrophoresis gels after PCR using primers and annealing temperatures for various locus tag targets as summarized in Table 1. No= PCR amplicons of appropriate size were not detected.

	1,2-DCA	1,2-DCP	1,2,3-TCP
% Dechlorination	31-53%	25-65%	41-49%
Incubation Time (Days)	8	21	24
Daughter Products	Ethene	Propene	Allyl Chloride
RIN Number	8.5-9.2	8.4-9.6	8.4-9.4
RNA concentration (ng/µL)	11.2-17.2	3.5-4.6	7.4-17.5
Dehly_0068	No	No	No
Dehly_0069	Yes	Yes	Yes
Dehly_0075	No	No	No
Dehly_0121	Yes	Yes	Yes
Dehly_0156- Dehly_0157	Yes	Yes	Yes
Dehly_0274	Yes	Yes	Yes
Dehly_0275- Dehly_0276	Yes	Yes	Yes
Dehly_0283	No	No	No
Dehly_0479	Yes	Yes	Yes
Dehly_0849	Yes	Yes	Yes
Dehly_0910	Yes	Yes	Yes
Dehly_1053-Dehly_1054	Yes	Yes	Yes
Dehly_1148	Yes	Yes	Yes
Dehly_1151-Dehly_1152	Yes	Yes	Yes
Dehly_1328	Yes	Yes	Yes
Dehly_1355	No	No	No
Dehly_1514	Yes	Yes	Yes
Dehly_1520	Yes	Yes	Yes
Dehly_1523	Yes	Yes	Yes
Dehly_1524-Dehly_1525	Yes	Yes	Yes
Dehly_1530	Yes	Yes	Yes
Dehly_1533-Dehly_1534	No	No	No
Dehly_1540	No	No	No
Dehly_1541	Yes	Yes	Yes
Dehly_1582	Yes	Yes	Yes

Simultaneous transcription of multiple *rdhA* genes has also been reported for mixed cultures that reductively dehalogenate chlorinated alkanes or alkenes (Tang & Edwards 2013; Tang *et al*., 2013). The results reported here for *D. lykanthroporepellens* further expand the diversity of bacteria known to simultaneously transcribe multiple *rdh* genes to include additional dehalorespiring species in the phylum *Chloroflexi*.

The expression pattern for *rdh* and regulatory genes was identical for BL-DC-9^T during dechlorination of all three chlorinated compounds, with transcripts detected for all replicates with primers targeting 19 of the *rdh* loci and transcripts not detected for any replicates using primers targeting six of the *rdh* loci (Table 3). This latter result suggests that genes with locus tags Dehly_0068, Dehly_0075, Dehly_0283, Dehly_1355, Dehly_1533/Dehly_1534, and Dehly_1540) are likely not directly involved in the dechlorination of the three tested substrates. They may be upregulated in presence of different halogenated substrates. It is also possible that these genes may have been transcribed but at levels below detection.

The gene with locus tag Dehly_1524 has been recently identified as encoding a 1,2-DCP reductive dehalogenase (DcpA) based on a combination of blue native polyacrylamide gel electrophoresis and enzyme assays (Padilla-Crespo *et al.*, 2014). Homologs of this gene (92% amino acid identity) have also been found in 1,2-DCP dechlorinating *Dehalococcoides mccartyi* cultures RC and KS (Padilla-Crespo *et al.*, 2014). In experiments reported here, the Dehly_1524 gene transcript was detected using RT-PCR with BL-DC-9^T cultures actively dechlorinating not only 1,2-DCP but also 1,2-DCA and 1,2,3-TCP. This leaves open the possibility that the same enzyme was responsible for transformation of all three halogenated alkanes.

Two types of transcriptional regulators, MarR and two-component system transcriptional regulators, were found to be associated with *rdh* genes in the genomes of *Dehalococcoides mccartyii* strains (Kube *et al*., 2005; Seshadri *et al*., 2005). MarR mediates bacterial response to changes in environmental stress and metabolic conditions, usually by functioning as a repressor protein (Wilkinson *et al*., 2006). The binding of a substrate to the MarR protein results in its release from the promoter, thereby de-repressing transcription (Providenti & Wyndham, 2001; Wagner *et al*., 2013). The two-component system transcriptional regulators, on the other hand,

contain a histidine kinase, which is phosphorylated on signal recognition. The phosphorylated signal is subsequently transferred to a response regulator containing a C-terminal DNA binding domain (Galperin, 2006).

The genome sequence shows the presence of a MarR orthologue located immediately upstream and in opposite orientation of RdhA Dehly_0121. Also at least three other putative RhdA proteins have two component system transcriptional regulators associated with them. These included Dehly_1148 that had Dehly_1149 (winged helix family two component system transcriptional regulator) with a cognate sensor signal transduction histidine kinase Dehly_1150 interspersed between and in the same orientation to that of Dehly_1148. Dehly_0270 was hypothesized to be the transcriptional regulator for Dehly_0274 as blastp indicates that it is a member of the LuxR family of two component transcriptional regulators. It is located upstream and in the same orientation to that of Dehly_0274 and it also has the cognate histidine kinase Dehly_0269. Dehly_1530 is also hypothesized to be regulated by Dehly_1531 (LuxR family two component system regulator). Dehly_1531 and its cognate histidine kinase are interspersed between and in the same orientation of Dehly_1530.

In the RT-PCR experiments reported here, the MarR homologue Dehly_0120 was detected along with transcripts from the associated *rdhA* (Dehly_0121) for all of the BL-DC-9^T cultures growing with halogenated solvents. Transcripts from three of the two-component system transcriptional regulators (Dehly_0270, Dehly_1149, and Dehly_1531) were also detected along with transcripts from their putatively associated *rdhA* genes (Dehly_0274, Dehly_1148 and Dehly_1530). This suggests that *rdhA* gene regulation is mediated by both types of transcription factors in *D. lykanthroporepellens* BL-DC-9^T.

2.5 Conclusion

The simultanous transcription of many *rdhA* genes and the expression of the same genes during the dechlorination of 1,2-DCA, 1,2-DCP, and 1,2,3-TCP argues against the direct implication of any one gene or group of genes with a set function. As with other genome sequences, the *rdhA* genes of *D. lykanthroporepellens* BL-DC-9^T were annotated on the basis of sequence identity with previously annotated genes, and their actual functionality remains to be elucidated. As noted previously, whether proteins encoded by the *rdhA* genes are even involved with reductive dechlorination remains tenuous (Hug *et al*., 2013). Nevertheless, identification of *rdhA* genes and transcription regulators expressed during the dechlorination of 1,2-DCA, 1,2-DCP, and 1,2,3-TCP sets the stage for future investigations into gene function and cell response to differing environmental conditions.

Chapter 3. Expression of 1,2-DCP reductive dehalogenase from *D. lykanthroporepellens* BL-DC-9^T in *E. coli*

3.1 Introduction

The experiments described in this chapter were aimed at heterologous expression of a 1,2-dichloropropane (1,2-DCP) reductive dehalogenase enzyme from *D. lykanthroporepellens* strain BL-DC-9^T using *E. coli* as the host. Initial experiments were also conducted to develop an appropriate dehalogenase enzyme assay for use in assessing the catalytic activity of the recombinant protein.

The reductive dehalogenase gene selected for cloning was the gene with locus tag Dehly_1524. As described in Chapter 2, this gene is one of the five "full length" *rdhA* genes with a cognate *rdhB* gene that was found to be transcribed during dechlorination of 1,2-DCA, 1,2-DCP, and 1,2,3-TCP. Additionally, Padilla-Crespo *et al*. (2014) recently implicated the Dehly_1524 gene as encoding an enzyme (designated as DcpA) that catalyzes a dihaloelimination reaction in which 1,2-DCP is transformed to the product propene. Their identification approach combined protein assays from blue native polyacrylamide gel electrophoresis (BN-PAGE), SDS-PAGE separation of the proteins eluted from the gel slice exhibiting 1,2-DCP dechlorination activity, and LC-MS/MS analysis of the proteins separated by SDS-PAGE (Padilla-Crespo *et al*., 2014). Dehly_1524 was the only reductive dehalogenase associated with the gel slices based on protein fragments identified.

The most closely related protein sequences to the DcpA of *D. lykanthroporepellens* BL-DC-9^T (92% amino acid identity) are two putative reductive dehalogenases (GenBank accession numbers JX826286 and JX826287) harbored by *D. mccartyi* strains KS and RC, respectively. These *Dehalococcoides* strains couple growth with 1,2-DCP-to-propene dechlorination in highly

enriched cultures (Padilla-Crespo *et al*., 2014). The use of cDNA generated from KS and RC cultures actively dechlorinating 1,2-DCP as template in qPCR with primers targeting the *dcpA* gene showed a significant upregulation of *dcpA* transcripts relative to starved controls. A graphical depiction of the organization of the *dcpA* and *dcpB* genes in *D. mccartyi* strains KS and RC is shown in Figure 3.1. The *D. lykanthroporepellens* BL-DC-9^T *dcpA* and *dcpB* genes (Dehly_1524 and Dehly_1525. respectively) have similar organization and encoded features.

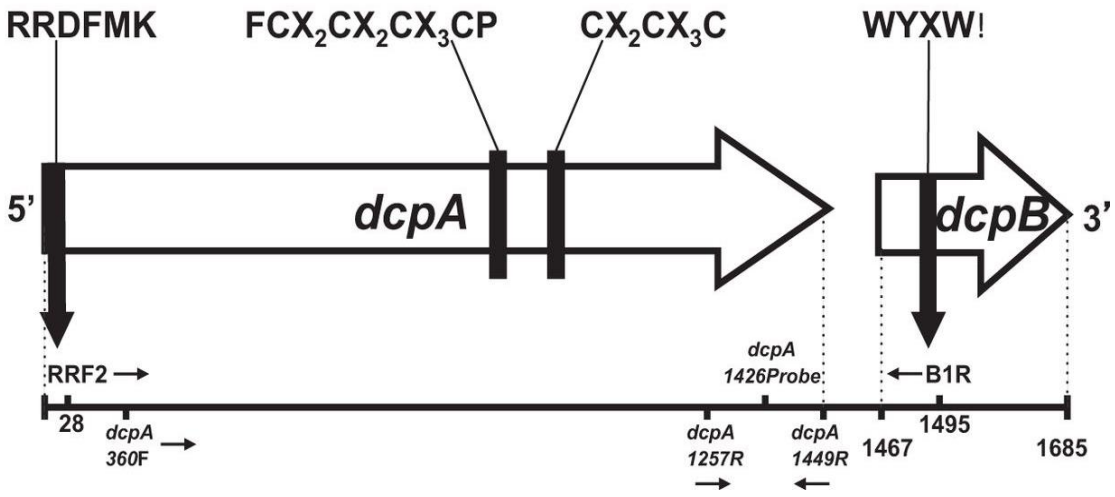


Figure 3.1 Arrangement of the *dcpA* and *dcpB* genes in *Dehalococcoides mccartyi* strains KS and RC (reprinted from Padilla Crespo *et al*., 2014). Characteristic features encoded by the *dcpA* gene, which include a Tat signal sequence (RRDFMK) near the N-terminus and two putative iron-sulfur cluster binding sequences containing 3 or 4 cysteine residues toward the C-terminus are noted at the top. Primer binding sequences are indicated at the bottom.

3.2 Materials and Methods

3.2.1 Synthesis of dehalogenase gene construct and subcloning in expression vector

The gene with locus tag Dehly_1524 was chemically synthesized by GenScript (Piscataway, NJ) with some modifications. First, the TAT signal sequence enabling transport of

putative membrane protein to the periplasmic space (Stanley *et al*., 2000) was not included in the gene construct. Second, the gene construct was codon optimized for the heterologous expression of the protein in *E. coli* by Jason Zhou at GenScript (NJ). The 5' flanking region was provided with the restriction site for NdeI 'CAT**ATG'** with the start codon in bold. Also, a BamHI site 'GGATCC' was included after the termination codon of the gene. The gene construct was subcloned in the BamHI and NdeI sites of the expression vector pET-16b (Novagen, Madison, WI, see Appendix D2 for plasmid map). The expression vector carries an isopropyl β-D-1-thiogalactopyranoside (IPTG) inducible promoter and a gene for ampicillin resistance. The gene was cloned in-frame with an N-terminal His Tag that is encoded by the expression vector. The subcloning in the pET-16b vector was verified by restriction digestion and gel electrophoresis at GenScript. The full-length, codon-optimized gene construct sequence and the original gene sequence are provided in Appendix D.

3.2.2 Development of dehalogenase enzyme assay using *Dehalogenimonas lykanthroporepellens* BL-DC-9^T cells

The dehalogenase enzyme assay was adapted from an assay that was originally developed by Holscher *et al*. (2003) and recently used by Tang *et al*. (2012). *D. lykanthroporepellens* BL-DC-9^T cells were grown in the presence of 0.5 mM 1,2-DCP as the electron acceptor using medium and incubation conditions as described by Maness *et al*. (2012). *D. lykanthroporepellens* BL-DC-9^T cells were collected by centrifugation of 400 mL of cell culture at 5,900×*g* at 4 °C for 30 minutes. Centrifugation was done in batches of 50 mL cell culture in plug sealed 50 mL centrifugation tubes (Corning). Working inside an anaerobic chamber (Coy) supplied with gas comprised of 10% H_2, 10% CO_2, and 80% N_2 (vol%), clarified supernatant was removed using a 25 mL pipette, leaving about 1 mL of culture in each centrifugation vial. The centrifugation vials

were vortexed to resuspend the pellets, and then the 1 mL re-suspended aliquots from eight centrifuge tubes were combined in a single 50 mL centrifugation tube. This was followed by another round of centrifugation at 5,900×*g* at 4 °C for 30 minutes. The supernatant was discarded after the centrifugation to obtain a cell pellet. Cell pellets were resuspended in 250 µL of 4× Sample prep Buffer (Native Gel Sample Prep Kit, Invitrogen, Cat. No. BN2008) containing 1% digitonin for detergent cell lysis. The resuspended cells were transfered to bead beating tubes from an UltraClean Microbial DNA Isolation Kit (MoBio Laboratories, Cat#12224-250-BT). The resuspended cells were lysed by mechanical sheering in a horizontal shaker (Vortex Genie 2 at 3,200 rpm) with two minutes of shaking followed by one minute of incubation on ice. This cycle was repeated three times. After this step the bead beating tubes were transferred inside the anaerobic chamber and for the enzyme assay, 20 µL of cell lysate was added to 3 mL of assay buffer [100 mM Tris HCl, 2 mM methyl viologen, and 0.5 mM 1,2-DCP] supplemented with titanium citrate solution [(Zehnder and Wuhrmann, 1976), 0.5 mL per 100 mL of assay buffer, resulting in a final concentration of 0.3 mM titanium and 0.9 mM citrate]. For comparison purposes, in the assay used by Tang *et al*. (2012), the final concentration of titanium citrate was 2 mM and a chlorinated substrate concentration of 30-70 mg/L was used. For preparation of the titanium citrate solution, 11.9 g of sodium citrate·$2H_2O$ was dissolved in 227 mL of water. The dissolved sodium citrate solution was transfered into an anaerobic chamber, and then 3.0 g of titanium chloride was added and mixed by swirling using a 25 mL serological pipette. The mixture was then distributed to 50 mL centrifuge tubes, taken out of the anaerobic chamber, and centrifuged at 8,000×*g* for 20 minutes. Following centrifugation, the supernatant was filter sterilized while working inside the anaerobic chamber and then it was stored in sealed sterile glass 160 mL serum bottles until use.

The enzyme assays were set up in 12 mL amber colored serum vials. After adding the cell lysate, the vials were sealed with butyl rubber stoppers and aluminum crimp caps and incubated inside the anaerobic chamber for 24 hours at 30 °C. Identical enzyme assays were also set up but without adding the cell lysate (negative control) and also using the supernatant (soluble fraction) collected after centrifugation of the cell lysate. After adding the cell lysate in the assay vials inside the anaerobic chamber the bead beating tubes were tightly capped and taken outside of the anaerobic chamber for centrifugation at 4000×*g* for two minutes. The bead beating tubes were taken back in the anaerobic chamber and 50 µl of the supernatant was also used in the enzyme assay to check if the enzyme activity was present both in the soluble and insoluble fractions. Following incubation, a 100 µL aliquot of gas headspace was collected from each serum bottle using a glass gas-tight syringe (Pressure-Lok, VICI Precision Sampling). Gas samples were then introduced by direct injection to a gas chromatograph (Agilent 7820A) equipped with a flame ionization detector (FID). Activity was assessed by measuring propene production and decrease in percentage of 1,2-DCP in bottles receiving cell lysates relative to negative controls.

3.2.3 Transformation of the expression vector into host cells

The codon-optimized gene construct Dehly_1524 cloned in the pET-16b expression vector (Novagen) was transformed in BL21(DE3)pLysS (Novagen) chemically competent *E. coli*. Cells were transformed by adding 1 µL of the plasmid solution (10 ng/µL concentration) to the solution containing BL21(DE3)pLysS cells. The mixture of cells and DNA was transferred to a water bath at 42 °C for 30 seconds and then shifted to an ice bath and incubated for 5 minutes. SOC media was added to the cells, and the cells were incubated at 37 °C for 60 minutes in a rotating shaker shaking at 250 rpm. The SOC media was used from BL21(DE3)pLysS singles kit

(Novagen Cat. No.70232-3). Following incubation, 25 and 50 µL aliquots of the cells were plated on LB agar plates containing 50 µg/mL ampicillin, 34 µg/mL chloramphenicol, and 1% glucose. The plates were incubated overnight at 37 °C and then visually inspected for the appearance of colonies.

3.2.4 Induction and overexpression of the protein

For a small scale protein induction, a single colony was picked and inoculated in 3 mL LB broth in a 10 mL falcon tube with 50 µg/mL ampicillin and 34 µg/mL chloramphenicol. The LB was also supplemented with 2× vitamin mix [4-aminobenzoic acid 40 mg, D(+)-Biotin 10 mg, nicotinic acid 100 mg, calcium D(+)-pantothenate 50 mg, pyridoxine dihydrochloride 150 mg dissolved in 500 mL of sodium phosphate buffer (10mM: pH 7.1)], 2× thiamine mix [100 mg thiamine chloride dihydrochloride dissolved in 500 mL of sodium phosphate buffer (25 mM; pH 3.4)], and 2× vitamin B_{12} solution (50 mg cyanocobalamine in 500 mL distilled water) (0.5 mL/L concentration for each).

To reach the exponential phase of cell growth, the cells were incubated overnight at 37 °C in a rotating shaker shaking at 250 rpm. The following day, 50 µL of the dense cell culture was inoculated into fresh 3 mL LB broth supplemented with the same amendments as before and incubated at 37 °C for 3 hours in a rotating shaker shaking at 250 rpm. After 3 hours, cultures were split into two parts each having 1.5 mL volume. One of these was induced by adding 3 µL of 0.5 M IPTG (isopropyl β-D-1-thiogalactopyronaside) solution resulting in a final IPTG concentration of 1 mM. The induced and non-induced cultures were further incubated at 37 °C in a rotating shaker at 250 rpm for two more hours. Cells from the non-induced and induced cell cultures were collected by centrifugation at 8,000×*g* for 1 minute. Cells were resuspended in 500 µL 4× Sample prep Buffer (Native Gel Sample Prep Kit, Invitrogen, Cat. No. BN2008)

containing 1% digitonin for detergent cell lysis. The resuspended cells were transferred to bead beating tubes from an UltraClean Microbial DNA Isolation Kit (MoBio Laboratories, Cat#12224-250-BT). The resuspended cells were lysed by mechanical sheering in a horizontal shaker (Vortex Genie 2 at 3,200 rpm) with four cycles each consisting of two minutes of shaking followed by one minute of incubation on ice. Enzyme assays were set up using the *E. coli* cell lysates the same way that were used for the *D. lykanthroporepellens* BL-DC-9^T cells with 1,2-DCP as substrate (see Section 3.2.2).

3.2.5 Induction experiments to obtain the protein in the soluble fraction and prevent protein misfolding

As described in the Results section of this chapter, the protein expressed by transformed *E. coli* cells under the conditions described above in Section 3.2.4 was not observed to have catalytic activity under the conditions tested, and the bulk of the recombinant protein was found to be associated with the insoluble fraction of cell lysates. In an attempt to increase the solubility of the protein and obtain the protein in a catalytically active form, experiments were performed using a variety of alternate induction conditions. These experiments included different induction temperatures, induction times, and IPTG concentrations as described in the following sub-sections.

3.2.5.1 Different induction temperatures

To explore whether induction at a temperature lower than the initially tested temperature of 37 °C would lead to production of a soluble and/or catalytically active dehalogenase, additional experiments were carried out with induction at 23 °C. Two tubes each containing 3.0 mL culture with antibiotics were inoculated from a single clone of transformed cells. After 2 hours of incubation at 37 °C, the culture volumes were each divided equally into two separate

tubes (resulting in 4 tubes total, each containing 1.5 mL culture) and 3 µL of 0.5 M IPTG was added to each of the four tubes and induced for 6 hours at 23 °C. After induction, the cells were collected by centrifugation at 8,000×*g* for 1 minute. The soluble and insoluble cell lysate were prepared using the BugBuster protein extraction reagent protocol (Novagen Cat. No.70584-3).The insoluble and the soluble fraction was analyzed by 12% SDS-PAGE (Mini-PROTEAN TGX Gels, BIO-RAD). The gels were run at 200 Volts for 45 minutes. 10× Tris/Glycine/SDS Buffer was used as the running buffer (BIO-RAD Cat. # 161-0732). Following electrophoresis, the gels were stained using 50 mL of Bio-Safe Coomassie G-250 stain (Biorad Cat. # 161-0786) for 45 minutes in a slow moving shaker. After staining, the gel was transferred to 50 mL deionized water and incubated in a slow moving shaker for about two hours till the bands started to appear.

In a separate experiment, induction was carried out at 18 °C. Two tubes of 3 mL culture with antibiotics were inoculated with a single clone of transformed cells. After 2 hours of incubation at 37 °C, the cultures were split apart and 3 µL of 0.5 M IPTG was added to each of the four tubes and induced for 10 hours at 18 °C. The longer incubation period for induction at 18°C (10 hours) was longer than that for induction at 23 °C (6 hours) and the initial base-line test at 37 °C (2 hours). After induction, the crude cell extract and the soluble fraction was analyzed by 12% SDS-PAGE as described above.

3.2.5.2 Different IPTG concentrations

To explore whether induction with an IPTG concentration different from the initial test (which involved addition of 3 µL of 0.5 M IPTG per 1.5 mL culture volume, final concentration 1 mM) would lead to production of a soluble and/or catalytically active dehalogenase, additional experiments were carried out with induction using lower IPTG concentrations. Two tubes of 3

mL culture with antibiotics were each inoculated with a single clone of transformed cells. After 2 hours of incubation at 37 °C, the cultures were divided and 3 µL aliquots of IPTG solution containing three different concentrations (0.1 M, 0.2 M, and 0.3 M) was added to the three tubes, which were induced for 2 hours at 37 °C (final IPTG concentrations ranging from 0.6 mM to 0.2 mM). After induction, the crude cell extract and the soluble fraction was analyzed by 12% SDS-PAGE.

3.2.5.3 Different induction times

To explore whether induction at a temperature of 37 °C but with a shorter induction time would lead to production of a soluble and/or catalytically active dehalogenase, additional experiments were carried out with alternate durations of incubation following IPTG addition prior to cell harvest. Two tubes of 3 mL culture with antibiotics were inoculated with a single clone of transformed cells. After 2 hours of incubation at 37 °C, the two cultures were each divided equally into separate tubes (four tubes total). Then, 3 µL 0.5 M IPTG was added to each of the four tubes and the cultures were induced for four different time durations (0.5, 1.0, 1.5, and 2.0 hours) at 37 °C. After induction, crude cell lysates and the soluble fractions were analyzed by 12% SDS-PAGE.

3.2.6 Protein refolding and enzyme assay

As described in the Results section of this chapter, the protein expressed by transformed *E. coli* cells under the conditions described above in Section 3.2.5 was not observed to have catalytic activity under the conditions tested. In an attempt to obtain the protein in a catalytically active form, experiments were performed to refold the protein from the insoluble fraction and

check the refolded protein for catalytic activity. Inculsion bodies were first purified from the insoluble fraction and refolding was attempted as described in the following subsections.

3.2.6.1 Preparation of inclusion bodies for protein refolding

100 mL of LB broth with antibiotics was innoculated with 1% v/v culture of transformed cells. After three hours of growth at 37 °C the cells were induced with IPTG resulting in a final concentration of 1mM. The induced cultures were further incubated at 37 °C in a rotating shaker at 250 rpm for two more hours. The induced cells were harvested by centrifugation at 10,000×*g* for 10 min using a weighed centrifuge tube. The harvested cell pellets were resuspended in 5 mL of BugBuster reagent (Novagen). After resuspension in BugBuster reagent, a majority of the steps were carried inside the anaerobic chamber to prevent oxygen exposure to the induced protein. The resuspended cells were incubated on a rotating shaker at a slow setting for 10-20 minutes at room temperature. Inclusion bodies were prepared following the BugBuster protein extraction reagent protocol (Novagen Cat. No.70584-3). After the incubation of the cell pellet with BugBuster reagent, the insoluble cell debris was removed by centrifugation at 16,000×*g* for 20 minutes at 4 °C. The pellet obtained after centrifugation was resuspended in the same volume of BugBuster reagent that was used to resuspend the original cell pellet. Resuspension was accomplished by complete vortexing. This was followed by an incubation of 0.5 hour, and then an equal volume of 1:10 diluted BugBuster reagent diluted in nuclease-free water was added to the suspension and mixed. The cell suspension was then centrifuged at 5,000×*g* for 15 min at 4 °C. Inclusion bodies were collected as the precipitate. The inclusion bodies were washed by two more rounds using 1:10 diluted BugBuster reagent diluted in nuclease-free water. The inclusion body suspension was analyzed by 12% SDS-PAGE, and the washed inclusion body pellets were stored at -80 °C until refolding.

3.2.6.2 Solubilization of inclusion bodies

The inclusion bodies were solubilized according to the recommended protocol from the Pierce™ Protein Refolding kit employed for subsequent refolding (see Section 3.2.6.3). Briefly, the inclusion body pellet was dissolved in 1 mL of 6-8 M guanidine hydrochloride, 50 mM Tris (pH 8.0). The suspension was vortexed to obtain a homogenized suspension. The suspension was incubated overnight at 4 °C to completely dissolve the inclusion body. After the overnight incubation, insoluble components were removed by centrifugation at 10,000×g for 20 minutes, and the supernatant was collected. The concentration of protein in the supernatant was quantified using a Qubit 2 protein assay kit (Life Technologies).

3.2.6.3 Refolding protocol

Protein refolding was attempted using a Pierce™ Protein Refolding Kit (Life Technologies, Cat. No. 89867). The refolding kit includes nine different buffers that provide a wide range of protein denaturing conditions. According to the manufacturer's protocol, the buffers provide nine different denaturation conditions that range from 0-1.4 M guanidine and 0-0.80 M arginine. Table 3.1 shows the matrix set up for the initial protein refolding screening experiment. Table 3.2 summarizes the secondary screening experiments for refolding in the presence of polyethylene glycol. The refolding reactions were set up inside the anaerobic chamber in 2 mL microcentrifuge tubes containing 900 µL aliquots of the base refolding buffers along with EDTA, reduced glutathione, oxidized glutathione, and water (Tables 3.1 and 3.2). The refolding buffer was also supplemented with 2× Vitamin mix, 2× Thiamine mix, and 2× Vitamin B_{12} solution (0.5 mL/L, concentration details about the vitamins provided in Section 3.2.4). The refolding tubes and the denatured protein solution was placed on ice. 10 µL of protein solution was added to each refolding tube, and after the addition, the tubes were vortexed

and immediately placed on ice. This was repeated until 50 µL of protein solution was added to each refolding tube. The refolding tubes were incubated for 24 hours at 4°C. The refolding tubes were transferred to 30 °C water bath for 2 hours, and then a fraction of the solution was analyzed for refolding using the dehalogenase enzyme assay described in Section 3.2.

Table 3.1 Tabular Representation of the Refolding Reaction tubes. First screening for refolded proteins in the presence of denaturing concentration, reducing environment and protein concentration.

Tube ID	Base Refolding Buffer ID No.	Volume 100 mM EDTA (µL)	Volume 200 mM GSH (µL)	Volume 100 mM GSSH (µL)	Volume H_2O (µL)	Volume 0.5-1.0 mg/mL protein stock (µL)
1	1	10	10	2	28	50
2	2	10	10	4	26	50
3	3	10	5	10	25	50
4	4	10	10	4	26	50
5	5	10	5	10	25	50
6	6	10	10	2	20	50
7	7	10	5	10	25	50
8	8	10	10	2	28	50
9	9	10	10	4	26	50

GSH-Reduced Glutathione

GSSH-Oxidized Glutathione

Table 3.2 Refolding screening in the presence of polyethylene glycol

Tube ID	Refolding Buffer using buffers with buffer ID nos 6 and 4 (900 µL)	Volume 100 mM EDTA (µL)	Volume 200 mM GSH (µL)	Volume 100 mM GSSG (µL)	Volume 10 mM PEG (µL)	Volume H_2O (µL)	Volume 0.5-1.0 mg/mL protein stock (µL)
1	6	10	10	2	0	28	50
2	6	10	10	2	6.7	21.3	50
3	6	10	10	2	13.3	14.7	50
4	6	10	10	2	20	8	50
5	4	10	10	2	0	28	50
6	4	10	10	2	6.7	21.3	50
7	4	10	10	2	13.3	14.7	50
8	4	10	10	2	20	8	50

PEG-Polyethylene Glycol (polymer that blocks aggregation during refolding)

GSH-Reduced Glutathione, GSSG-Oxidized Glutathione

3.2.7 Experiments to adress potential reductive dehalogenase ligand binding issues

As described in the Results section of this chapter, the protein refolding described above in Section 3.2.6 was not observed to result in catalytic activity of the recombinant protein under the conditions tested. In an attempt to obtain the protein in a catalytically active form, further experiments were performed to address potential reductive dehalogenase ligand binding issues by mixing the cells lysates of the induced *E. coli* cells with cell lysates from *Dehalogenimonas alkenigignens* IP3-3^T cells.

400 mL of *D. alkenigignens* IP3-3^T cells actively dechlorinating 1,2-DCP were collected by centrifugation at 5900×*g* at 4 °C for 30 minutes. The centrifugation was done in batches of 50 mL cell culture in plug sealed 50 mL centrifugation tube (Corning). The supernatant was discarded using a 25 mL pipette leaving about 1 mL of culture in the centrifugation vial inside the anaerobic chamber. Approximately 8 mL of culture was collected after the first centrifugation step, and the culture was centrifuged again at 3,500 rpm (5,900×*g*) at 4 °C to obtain the cell pellet. The cell pellets were resuspended in 250 µL of 4× Sample prep Buffer (Native Gel Sample Prep Kit, Invitrogen, Cat. No. BN2008) containing 1% digitonin for detergent cell lysis. The resuspended cells were lysed by mechanical sheering in a horizontal shaker. The cell lysis was accomplished by two minutes of shaking followed by one minute of incubation on ice. At the same time, a medium scale protein induction was set up using 50 mL of induced cell culture. 10 µL of 0.5 M IPTG was used for inducing 50 mL of transformed *E. coli* cell culture having Dehly_1524 cloned in pET-16b expression vector (Novagen). The induction was carried out at 37 °C in a rotating shaker at 250 rpm for two more hours. The induced cells were harvested by centrifugation at 10,000×*g* for 10 minutes using a weighed centrifuge tube. The harvested cell pellets were resuspended in 5 mL of BugBuster reagent (Novagen). The

resuspended cells were incubated on a rotating shaker at a slow setting for 10-20 minutes at room temperature. The crude cell lysates of *D. alkenigignens* IP3-3^T and the *E. coli* cells were mixed in the ratio 1:1, and 50 µL of the mixture was used for the dehalogenase enzyme assay described in Section 3.2.2. As a control, 50 µL of crude cell lysate of *D. alkenigignens* IP3-3^T cells were also used in the dehalogenase enzyme assay.

3.3 Results

3.3.1 Enzyme assay results with *D. lykanthroporepellens* BL-DC-9^T cell lysate

Triplicate experiments were set up, and dechlorination activity was monitored after overnight incubation of the assay vials by headspace GC analysis. 10-15% dechlorination was observed consistently after overnight incubation in vials with the crude cell lysate added. No dechlorination was observed in assay vials with the cell supernatant, indicating enzymatic activity to be present in the insoluble cell fraction. Negative control assay vials that did not receive part of the cell lysate also did not exhibit dechlorination. The assay developed was subsequently used in the detection of dehalogenase activity of the induced dehalogenase protein in recombinant *E. coli*.

3.3.2 Transformation results

The LB agar plates having 50 µg/mL ampicillin, 34 µg/mL chloramphenicol, and 1% (m/v) glucose were incubated overnight and then visually inspected for the appearance of colonies. Transformed BL21(DE3)pLysS cells were observed the following day on the selection plates.

3.3.3 Expression of the 1,2-DCP reductive dehalogenase protein from *D. lyknathroporepellens* BL-DC-9^T in *E. coli*

In the initial experiment (induction with IPTG at a final concentration of 1 mM at 37°C for 2 hours) to check for the expression of the induced DcpA protein in tranformed *E. coli*, a fraction of the induced and uninduced cell culture was analyzed by 12% SDS-PAGE. The bulk of the expected sized protein fragment of 53 kDa was detected in the insoluble fraction of cell lysates (Figure 3.3).

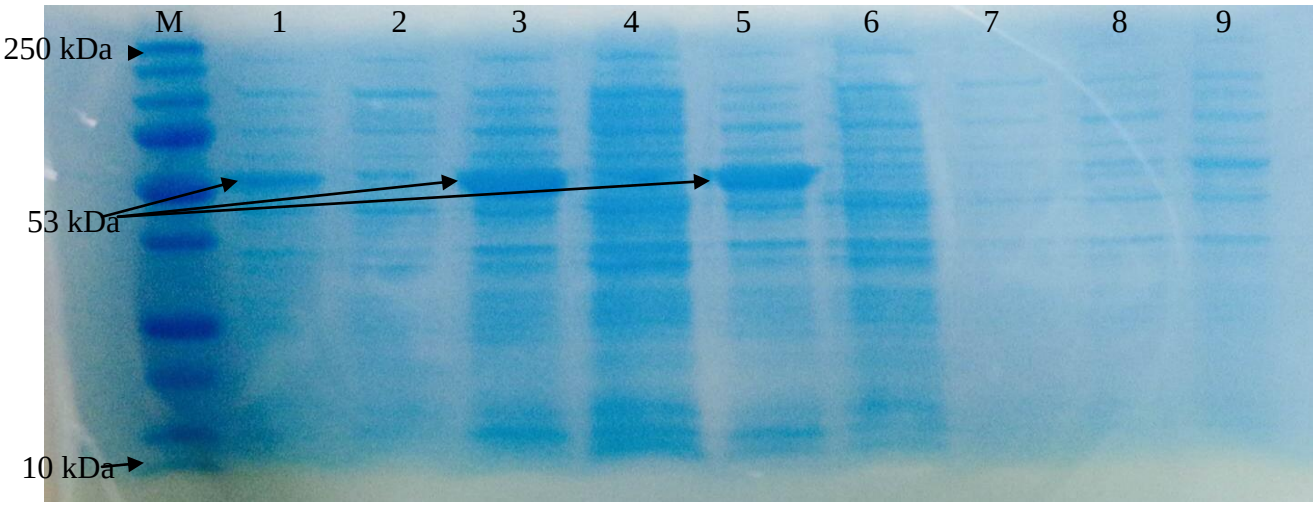


Figure 3.2 SDS-PAGE analysis of the transformed *E. coli*. Lanes 1, 3 and 5: cell lysate generated after after induction with 1 mM IPTG for two hours. Lanes 2, 4, and 6: cell lysate generated from uninduced cells grown under identical conditions. Lanes 7, 8, and 9: soluble fraction of cell lysate generated after after induction with 1 mM IPTG for two hours. 25 µL of cell extract was mixed with 25 µL of loading buffer, and the resulting 50 µL aliquot was loaded for lanes 1-9. Different lanes represent separate replicates. Lane M: 10 µL of Precision Plus Prestained All Blue Protein Standard (Biorad).

3.3.4 Enzyme assay results from the induced proteins

For initial "small scale (10 mL cultures)" enzyme assays using different volumes of induced *E. coli* cells in 10 mL total volume induced at a final IPTG concentration of 1 mM for 2

hours at 37°C, no dechlorination activity was observed. To investigate whether a large culture volume would yield activity, a "medium scale (50 mL culture)" protein induction was set up using 50 mL of induced cell culture with the same induction conditions; however, no enzyme activity was observed from 50 mL of induced cells. To check whether induction was successful, a fraction of the induced cell culture was analyzed by 12% SDS-PAGE (shown in Figure 3.3). The expected sized protein was observed in the SDS-PAGE analysis in the insoluble fraction. Similarly, a "large scale (200 mL and 500 mL culture)" induction was also set up using 200 mL and 500 mL of induced cell culture (again with the same induction conditions), and still enzyme activity was not observed using the dehalogenase enzyme assay. 10-15% dechlorination was consistently observed in positive controls with cell lysates from *D.lykanthroporepellens* BL-DC-9^{T}.

Using alternate induction conditions as described in Sections 3.2.5.1 to 3.3.5.3 had little effect in increasing the protein solubility, and the majority of the recombinant protein was consistently detected in the insoluble fraction. Expression of the protein was detected primarily in the insoluble fraction even when the expression was attempted at lower temperatues of 23 °C at a final IPTG concentration of 1 mM (see Figure D1 in Appendix D). When induction was attempted at shorter induction times of 0.5 hours, 1 hours, and 1.5 hours at 37 °C at a final IPTG concentration of 1 mM, it was consistently found to be expressed in the insoluble fraction in bulk (see Figure D2 in Appendix D). Expression of the protein was again observed mainly in the insoluble fraction when induction was attempted at lower final IPTG concentrations of 0.2 mM to 0.6 mM at 37 °C (see Figure D3 in Appendix D).

3.3.5 Enzyme assay results from the protein refolding and ligand binding experiments

After the enzyme assays from the "large scale" induction experiments (employing 500 mL culture volume with induction at the base-line conditions of 1mM IPTG did not show dehalogenase activity, additional experiments were undertaken in an attempt to refold the potentially mis-folded protein. After the inclusion bodies were prepared using the Inclusion Body Preparation protocol from the BugBuster protein extraction reagent (Novagen) a small fraction of the purified inclusion bodies were analyzed using 12% SDS-PAGE. The remaining inclusion body was solubilized using 1 mL of 6-8 M Guanidine Hydrochloride, 50 mM Tris (pH 8.0). The concentration of protein that was present in the solubilized inclusion body was found to be 3,000 ng/µL as estimated by Qubit assay (Life Technologies). Dehalogenase enzyme activity was not observed from any of the refolded protein solution tubes.

For the dehalogenase enzyme assay from the "large scale" induction experiments employing base-line conditions (500 mL culture volume, induction at 1 mM final IPTG concentration], 750 µg of protein (crude lysate) was applied in the assay. The Mac Nelly *et al.* (2014) group also used a similar range of 250-750 µg of protein (crude lysate) in the dehalogenase assay that was successful in detecting PceA dehalogenase activity after protein expression in *S. blattae*. In the assays from the refolding experiments reported in this chapter, however, 50 µg of protein was applied in the assay. In the most recent success at producing a functional dehalogenase using *E. coli*, a very high concentration of purified and cofactor-reconstituted vinyl chloride reductive dehalogenase was tested; a concentration of 0.2-1.0 mg/mL (with 6 mL per assay bottle) was shown to exhibit dehalogenase activity (Parthasarathy *et al*., 2015).

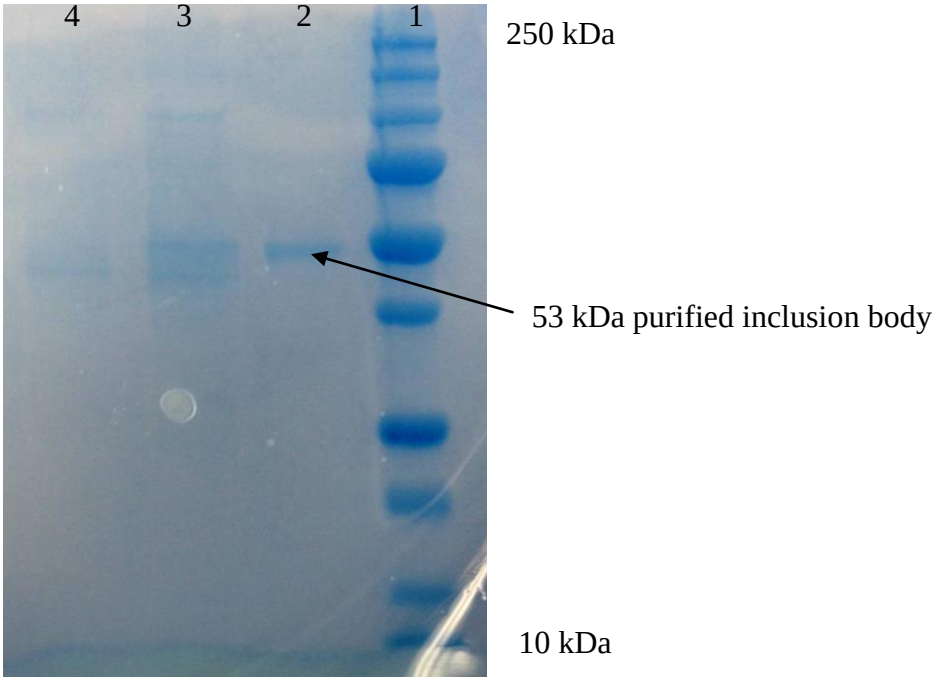


Figure 3.3 SDS-PAGE analysis after inclusion body purification. Predominantly a prominent band of expected size is seen in the lane 2 with purified inclusion body. Lane 3 is the induced *E. coli* cell lysate induced at 37 °C using a final IPTG concentration of 1 mM, and lane 4 is the uninduced cell lysate grown at 37 °C. Lane 1 is the Precision plus Protein standard (Biorad).

For the enzyme assays using mixtures of induced lysates of *E. coli* and *D. alkenigignens* IP3-3^T, approximately 50% or less enzyme activity was observed in the dehalogenase enzyme assay when the lysates of *E. coli* and *D. alkenigignens* IP3-3^T were mixed in the ratio 1:1 in comparison to the positive controls using only BL-DC-9^T cell lysate.

3.4 Discussion

At the time when functional expression of the 1,2-DCP reductive dehalogenase protein was attempted in experiments described in this dissertation chapter, a genetic system for the functional expression of dehalogenase enzymes was unknown. Hence, the expression was attempted using the standard host *E. coli*. Extreme oxygen sensitivity of this family of proteins makes the characterization of the dehalogenase protein challenging (Mac Nelly *et al*., 2014). The

gene was synthesized without the TAT signal sequence based on the rationale that it would be preferable to locate the functional protein inside the cytoplasm of the *E. coli* cells and thereby prevent it from reaching the oxygen rich periplasmic space of *E. coli*. The gene was codon optimized for expression in *E. coli* cells. The 53 kDa 1,2-DCP reductive dehalogenase protein was found to be expressed consistently in *E. coli* cells. In all attempts described in this chapter, however, the bulk of the protein was expressed in the insoluble fraction, and it was found to be catalytically inactive. The protein was consistently expressed in the insoluble fraction in spite of using lower induction times and lower IPTG concentrations for the induction. Solubility was not appreciably enhanced and catalytic activity was not detected when induction was carried out at lower temperature of 23 °C and 18 °C.

It is possible that the TAT signal sequence is involved in protein folding, and removing the TAT signal sequence may have resulted in the protein misfolding causing the protein to be expressed in the insoluble fraction. Enzyme activity for positive controls (*D. lykanthroporepellens* BL-DC-9^{T}) were also observed only with the insoluble fractions of crude cell lysates; enzyme activity was absent when the cell lysate was centrifuged and the supernatant was used in the enzyme assays. The protein may simply not be soluble. Also, in spite of using cobamides in the LB media during expression of the protein, the protein was not expressed functionally. All attempts to refold the protein from the insoluble fraction did not produce a catalytically active protein.

In previously reported attempts of expression of a functional dehalogenase in *E. coli*, the absence of a *de novo* cobamide biosynthesis pathway was speculated to be a likely reason for the failure to produce a functional protein (Suyama *et al*., 2002, Sjuts *et al*., 2012. This is further supported by the fact that the first report of a functional dehalogenase production heterologously

was in the cobamide producing *Shimwellia blattae* cells (Mac Nelly *et al*., 2014). This study reported that both 5,6-dimethylbenzimidazole and adeninyl-cobamides function as lower ligand cofactors for the PceA reductive dehalogenase and a certain level of endogenous cobamide is essential for the production of a functional dehalogenase enzyme. Cobamide extraction and analysis in this study showed that the *S. blattae* cells could produce pseudo-B_{12} that has dimethylbenzimidazole or adeninyl as the lower ligand (Mac Nelly *et al*., 2014). Also, when externally 5,6-dimethylbenzimidazole was added it possibly replaced adeninyl as the lower ligand and thereby increased cobamide production and the enzyme activity. It has been very recently reported that the crystal structure of the NprdhA reductive dehalogenase indicates a mechanism of cobamide dependent dehalogenation (Payne *et al*., 2015). Functional expression of the NprdhA protein was not detected in *E. coli,* again possibly due to the lack of cobalamin (Payne *et al*., 2015). The NprdhA lacks the TAT signal sequence indicating that it is a cytoplasmic protein and this was expressed successfully in the soluble fraction in *E. coli,* however, catalytic activity was not observed (Payne *et al*., 2015). This indicates that presence of the cobalamin cofactor is necessary for functional activity even when the protein is not misfolded and expressed in the soluble fraction. The first chapter of the dissertation mentions the significance of the cobamides in the dehalogenation chemistry. Reduction of the neutral cobalt(II) of the cobalamin to the superactive cobalt(I) form is likely to be an important step in the dehalogenation mechanism, and the lower ligand at the active site of the enzyme may carry out this conversion (Bommer *et al*., 2014). Since the *E. coli* cells lack a pathway for catabolic metabolism (Lawrence *et al*., 1996) that are cobamide dependent it can be concluded that production of a nonfunctional enzyme in *E. coli* is possibly due to to the lack of the cobamide synthesis pathway. The expression of the non-functional 1,2-DCP reductive dehalogenase

protein in BL21(DE3)pLysS *E. coli* cells is consistent with the notion that a *de novo* cobamide biosynthesis pathway is essential for maturation (catalytic activity) of the dehalogenase enzyme.

All characterized reductive dehalogenases have cobamide binding sites. The cobamides are members of vitamin B12 family that are produced in prokaryotes, and these cobamides play important role in regulation of gene expression in prokaryotes (Hazra *et al*., 2013). The main differences in the cobamides produced by the different organisms are in the structure of the lower axial ligand. To investigate whether the inactive form of the enzyme might be due to a difference in ligand binding, cell lysates of *E. coli* and *D. alkenigignens* IP3-3^T were mixed in the ratio 1:1 (v/v) and subjected to the dehalogenase enzyme activity assay. Dehalogenase enzyme activity in the mixture of induced *E. coli* and *D. alkenigignens* IP3-3^T was approximately 50% less in comparison to the positive controls, indicating that mixing the *E. coli* and *D. alkenigignens* IP3-3^T lysates did not produce a functionally active enzyme. This could be because the cobamide cofactor must be available during protein synthesis and folding. Also it is a possibility that cobamide production in *D. alkenigignens* IP3-3^T is different from the cobamide usage needed for the functional activity of DcpA in *D. lykanthroporepellenss* BL-DC-9^T.

3.5 Conclusions

The codon optimized 1,2-DCP reductive dehalogenase protein from *D. lykanthroporepellens* strain BL-DC-9^T was non-functionally expressed using *E. coli* as the host. The majority of the recombinant protein was expressed in the insoluble fraction. Dehalogenation activity was also associated with the insoluble fraction of *D. lykanthroporepellens* strain BL-DC-9^T, suggesting that the protein is poorly soluble, even if assembled with the appropriate cofactors. Attempts to refold the protein and mixing the *E. coli* and *D. alkenigignens* IP3-3^T cell lysate did not show an increase in functional enzyme activity, perhaps because the cobamides available

differed from the cobamide variant synthesized by *D. lykanthroporepellens* strain BL-DC-9^{T}. Based on the now well-documented requirement for a cobamide cofactor for dehalogenation activity, the standard genetic system of *E. coli* is unlikely to be a suitable host for the functional production of this dehalogenase.

Chapter 4. Expression of 1,2-DCP reductive dehalogenase from *Dehalogenimonas lykanthroporepellens* BL-DC-9^T in *S. blattae* DSM 4481^T

4.1 Introduction

As described in Chapter 1 (see Section 1.6), *Shimwellia blattae* DSM 4481^T has been recently reported to contain a genetic system that was successful in heterologous expression of catalytically active RdhA proteins from *Desulfitobacterium hafniense* strain Y51 and *D. hafniense* DCB-2. The tetrachloethene reductive dehalogenase (PceA) of the *Desulfitobacterium hafniense* strain Y51 was functionally expressed in the *S. blattae* cells. The active RdhA3 from *D. hafniense* DCB-2 was also expressed in *Shimwellia blattae* DSM 4481^T and the recombinant enzyme exhibited the conversion of 3,5-dichlorophenol,2,3-dichlorophenol and 2,4-dichlorophenol (Mac Nelly *et al*., 2014). The experiments described in this chapter were aimed at extending on this previous finding with the functional expression of a 1,2-DCP reductive dehalogenase enzyme from *D. lykanthroporepellens* strain BL-DC-9^T using *S. blattae* DSM 4481^T as the host.

4.2 Materials and Methods

4.2.1 Synthesis of dehalogenase gene construct and subcloning in expression vector

The gene sequence with locus tag Dehly_1524 was chemically synthesized at GenScript (NJ). In contrast with the approach employed in experiments described in Chapter 3, the TAT signal sequence essential for the transport of membrane proteins to the periplasmic space was included in the gene construct. This was based on the rationale that three previous studies have shown that it might be essential for proper folding of the Rdh protein (Sjuts *et al*., 2012; Malliard *et al*., 2011; Mac Nelly *et al*., 2014). The 5' flanking region was provided with the restriction site

for EcoRI 'GAATTC'. Also, a BamHI site 'GGATCC' was included after the termination codon of the gene. Codon optimization was not performed for expression in *S. blattae* as the functional PceA from *Desulfitobacterium hafniense* strain Y51 was expressed in *S. blattae* without codon optimization. The gene construct was subcloned in the EcoRI and BamHI site of the expression vector pASK-IBA 37 Plus (IBA-Germany, See Appendix E2 for plasmid map and sequence of the cloned gene). The gene was cloned in frame with an N-terminal His Tag that is encoded by the expression vector. The subcloning in the expression vector was verified by restriction digestion and gel electrophoresis at GenScript. The expression vector carries a tetracycline inducible promoter for the expression of cloned target proteins and a gene for ampicillin resistence.

For co-expression along with the DcpA protein. the *PceT* gene from *Desulfitobacterium hafniense* strain Y51 (trigger factor that has shown to increase the solubility of PceA protein in *S. blattae* cells, Mac Nelly *et al*., 2014) (GenBank Accession AP008230 region "complement 3227661-3228617") was chemically synthesized at Genscript (Piscataway, NJ) and subcloned in the EcoRI and BamHI cloning sites of of the expression vector pASK-IBA 3C (IBA-Germany, See Appendix E2 for vector map). The gene was cloned in-frame with a C-terminal Strep Tag that was encoded by the expression vector. The full-length original gene sequence is provided in appendix E2. The subcloning in the pASK-IBA 3C vector was verified by restriction digestion and gel electrophoresis at GenScript. The full-length, original gene sequence is provided in Appendix E2. This expression vector carries an anhydrotetracycline inducible promoter and has chloramphenicol resistance.

4.2.2 Preparation of competent *S. blattae* cells

S. blattae strain DSM 4481^T was purchased from the Deutsche Sammlung von Mikroorganismen un Zallkulturen GmbH (Braunschweig, Germany) in frozen form. Frozen *S. blattae* cells were reconstituted by adding liquid Nutrient broth (Difco) and 10% glycerol stocks of *S. blattae* DSM 4481^T cells were prepared immediately for long term storage at -80 ºC. For culture maintenance, *S. blattae* DSM 4481^T was cultivated aerobically on LB agar plates [(LB Broth Lennox (Sigma) supplemented with 20 g/L agar (Difco)] at 37°C.

For preparation of competent *S. blattae* cells, a single colony of *S. blattae* cells from an LB agar plate was innoculated into 10 mL of LB and grown overnight at 37°C. Then, 1.0 liter of LB was innoculated with 1% (v/v) overnight culture and grown until A600 = 0.5 to 1.0. The cells were subsequently chilled on ice for 15-30 minutes and centrifuged at 4.000×*g* for 15 min. Following centrifugation, supernatant was discarded, and the pellet was resuspended in 1 liter ice cold sterile water. That was followed by two more rounds of centrifugation at 4.000×*g* for 15 min and washing the pellets in ice cold water. The cells were then resuspended in 20 mL ice cold sterile 10% glycerol prepared in deionized water and centrifuged at 4.000×*g* for 15 min. The pellets were then resuspended in 2-3 mL of 10% glycerol, and 40 µL aliquots of the cells were flash frozen using liquid nitrogen and stored at -80 °C until use in cloning.

4.2.3 Transformation of the expression vector into host cells

40 µL of competent *S. blattae* cells prepared as described above were mixed with 8 ng of recombinant plasmid DNA (pASK-IBA 37 plus/Dehly_1524) and then transferred to the bottom of a cuvette that was pre-chilled at -20 °C. The mixture was subjected to electroporation (Biorad electroporator) and immediately 1 mL of SOC was added. The SOC media was used from BL21(DE3)pLysS singles kit (Novagen Cat. No.70232-3). The mixture of cells was grown for

one hour at 37 °C, and then 50 µL aliquots were plated on LB agar plates supplemented with 50 µg/mL ampicillin. The plates were incubated overnight aerobically at a temperature of 37°C before visual inspection to detect the appearance of colonies.

When transformation with two plasmids was attempted, the two plasmids were mixed and 8 ng mixture of recombinant plasmids (pASK-IBA 37 plus/Dehly_1524) and (pASK-IBA-3C/PceT) was added to 40 µL of competent *S. blattae* cells. The mixture was subjected to electroporation and after adding SOC media, cells were grown for one hour at 37 °C. Transformed cells were selected on LB agar plates supplemented with 50 µg/mL ampicillin and 34 µg/mL chloramphenicol.

4.2.4 Induction and overexpression of the Dehly_1524 protein under anoxic conditions

For expression of the recombinant reductive dehalogenase protein in transformed *S. blattae*, cells were grown in a mineral medium containing (per liter): K_2HPO_4, (14 g) KH_2PO_4, (6 g) $(NH_4)_2SO_4$, (3 g) $MgSO_4 \cdot 7H_20$ (0.02 g), $CoCl_2 \cdot 6H_2O$ (0.024 g), yeast extract (2 g), cysteine-HCl (0.2 g), trace element solution 1 mL (see Appendix E2 for composition) and glycerol (300 mM final concentration) (ph 7.5) as growth substrate along with 100 µg/mL ampicillin. Glycerol served as the carbon and energy source for the cells for cobamide production. External cobamide sources such as 5,6-dimethylbenzimidazole (DMB) (Sigma Cat. No. D147206-25G) and hydroxocobalmin (Sigma Cat. No. 95200-250mg) were added to the growth medium. Recent data has shown that the addition of these two external sources has a positive effect on the reductive dehalogenase activity of the protein (Mac Nelly *et al*., 2014). The final concentrations of dimethylbenzimidazole (DMB) and hydroxocobalmin were 10 µM and 10 nM, respectively.

For the expression of the reductive dehalogenase protein, *S. blattae* cells harboring the recombinant plasmid pASK-IBA 37 plus/Dehly_1524 were grown in LB broth for 6 to 8 hours at

28 °C. This culture was used to innoculate 250 mL of anaerobic mineral medium contained in 400 mL glass serum bottles. After dispensing the media in the bottles, the bottles were sealed using butyl rubber septa and aluminium crimp caps. The headspace was changed by flushing with filter sterilized Nitrogen gas. After flushing the headspace with nitrogen, the bottles were autoclaved, and filter sterilized ampicillin 100 µg/mL and vitamins dimethylbenzimidazole (DMB) and hydroxocobalmin were 10 µM and 10 nM, were added before inoculating with the transformed cells. When the PceT co expression was attempted, 34 µg/mL of chloramphenicol was also added in the anaerobic mineral medium. This anaerobic culture was grown at 18 °C with shaking at 100 rpm. After 6 hours when the OD reached 0.2-0.3, the cells were induced for protein expression by the addition of anhydrotetracycline (20 ng/mL final concentration). After 7-8 hours of induction, the cells were collected by centrifugation 5,900×*g* for 10 minutes at 4 °C and then transfered into the anaerobic chamber. The cells were resuspended in anoxic buffer (10 mM Tris-HCl made anoxic by flushing with filter sterilized N_2 gas), and cell lysis was accomplished by detergent lysis using glass beads and digitonin (detergent). As an alternative method, the cells were also lysed using BugBuster protein extraction reagent (Novagen) with cell lysate prepared following the protocol provided by Novagen. The BugBuster protein extraction reagent was made anaerobic by flushing with filter sterilized Nitrogen gas and the addition of titanium citrate [0.5 mL of prepared titanium citrate (details provided in Section 3.2.2) per 10 mL of BugBuster reagent]. The cell extracts were then used for the dehalogenase assay developed previously using 400 mL culture of wild type *D. lykanthroporepellens* BL-DC-9^T cells. Uninduced *S. blattae* cells were used as a negative control in the enzyme assay. *D. lykanthroporepellens* BL-DC-9^T cells grown with 0.5 mM 1,2-DCP as the electron acceptor were used as positive control in the enzyme assay.

4.2.5 Detection of the expressed dehalogenase enzyme

After 7-8 hours of induction, 500 µL of cell culture was extracted from the serum bottle using a syringe and cells were lysed using the BugBuster protein extraction reagent. A fraction of the suluble and insoluble fraction of the cell lysate were analyzed by 12% SDS-PAGE to check for the presence of the induced protein in the induced and uninduced cell culture.

4.2.6 Experiments attempted to express the protein in the soluble fraction

As described in the Results section of this chapter, the protein expressed by transformed *S. blattae* cells under the conditions described above in Section 4.2.4 was not observed to have catalytic activity under the conditions tested, and it was found to be associated with the insoluble fraction of cell lysates. In an attempt to obtain the protein in a soluble and catalytically active form, experiments were performed using a variety of alternate induction conditions. Induction was carried out at lower anhydrotetracycline concentrations and at shorter induction times. Two different culture volumes of 250 mL were innoculated with a single clone of transformed cells, and when the OD reached 0.25, induction was carried out at the final anhydrotetracycline concentrations of 10 ng/mL and 20 ng/mL. As a control, 250 mL of uninduced cell culture were also grown. 500 µL of induced cells were collected using a syringe at time intervals of 3, 5, and 7 hours, and then analyzed for expression of the protein in the soluble fraction by running SDS-PAGE.

4.2.7 Experiments to address potential reductive dehalogenase ligand binding issues

As described in the Results section of this chapter, the protein expression described above in Section 4.2.6 and 4.2.4 was not observed to have catalytic activity under the conditions tested. In an attempt to obtain the protein in catalytically active form, experiments were performed to

address the potential reductive dehalogenase ligand binding issue by mixing the cell lysates of the induced *S. blattae* cells and *D. lykanthroporepellens* BL-DC-9^T cells.

400 mL of *D. lykanthroporepellens* BL-DC-9^T cells actively dechlorinating 1,2-DCP were collected by centrifugation at 5,900×*g* at 4 °C for 30 minutes. The centrifugation was done in batches of 50 mL cell culture in plug sealed 50 mL centrifugation tube (Corning). The supernatant was discarded using a 25 mL pipette leaving about 1 mL of culture in the centrifugation vial inside the anaerobic chamber. Approximately 8 mL of culture was collected after the first centrifugation step and the culture was centrifuged again at 5,900×*g* at 4 °C to obtain the cell pellet. The cell pellets were resuspended in 250 µL of 4× Sample prep Buffer (Native Gel Sample Prep Kit, Invitrogen, Cat. No. BN2008) containing 1% digitonin for detergent cell lysis. The resuspended cells were lysed by mechanical sheering in a horizontal shaker. The cell lysis was accomplished by two minutes of shaking followed by one minute of incubation on ice. At the same time, 250 mL of induced cell culture of transformed *S. blattae* cells were collected by centrifugation without exposing the cells to oxygen and cell pellets were resuspended in 5 mL of BugBuster reagent (Novagen) inside the anaerobic chamber. The resuspended cells were incubated on a rotating shaker at a slow setting for 10-20 minutes at room temperature. The cell lysates of *D. lykanthroporepellens* BL-DC-9^T and the transformed *S. blattae* cells cells were mixed in the ratio 1:1, and 50 µL of the mixture was used for the dehalogenase enzyme assay. As a positive control, 50 µL aliquots of crude cell lysate of *D. lykanthroporepellens* BL-DC-9^T cells were also used in the dehalogenase enzyme assay.

4.3 Results

4.3.1 Transformation results

The LB agar plates having 50 µg/mL ampicillin were incubated overnight for the appearance of the transformed colonies. Transformed *S. blattae* cells with only the pASK-IBA 37 plus/Dehly_1524 (ampicillin resistant) were observed the following day on the selection plates.

4.3.2 Expression of the 1,2-DCP reductive dehalogenase protein from strain BL-DC-9^T in *S. blattae*

To check for the expression of the induced 1,2-DCP protein in transformed *S. blattae* cells containing pASK-IBA 37 plus/Dehly_1524, a fraction of the induced and uninduced cell culture was analyzed by 12% SDS-PAGE as described previously in Section 3.2.5.1. The expected sized protein fragment of 53 kDa was detected consistently predominently in the insoluble fraction of the induced cultures (Figure 4.2). The expression of the protein was consistently observed after induction for 8 hours at 18 °C at a final anhydrotetracycline concentration of 20 ng/µL.

Recombinant *S. blattae* cells that were selected with two antibiotics were also induced by anhydrotetracycline (final concentration 20 ng/mL) and induced for 8 hours at 18 °C. However the PceT co-expression was not seen when cell lysates were analyzed by 12% SDS-PAGE (see Appendix E1).

Regardless of induction conditions, the protein was consistently detected mostly in the insoluble fraction. Expression of the protein was detected in the insoluble fraction when the expression was attempted at lower anhydrotetracycline concentration of 10 ng/µL and after induction for 3, 5, and 7 hours respectively (see Figure E2 in Appendix E).

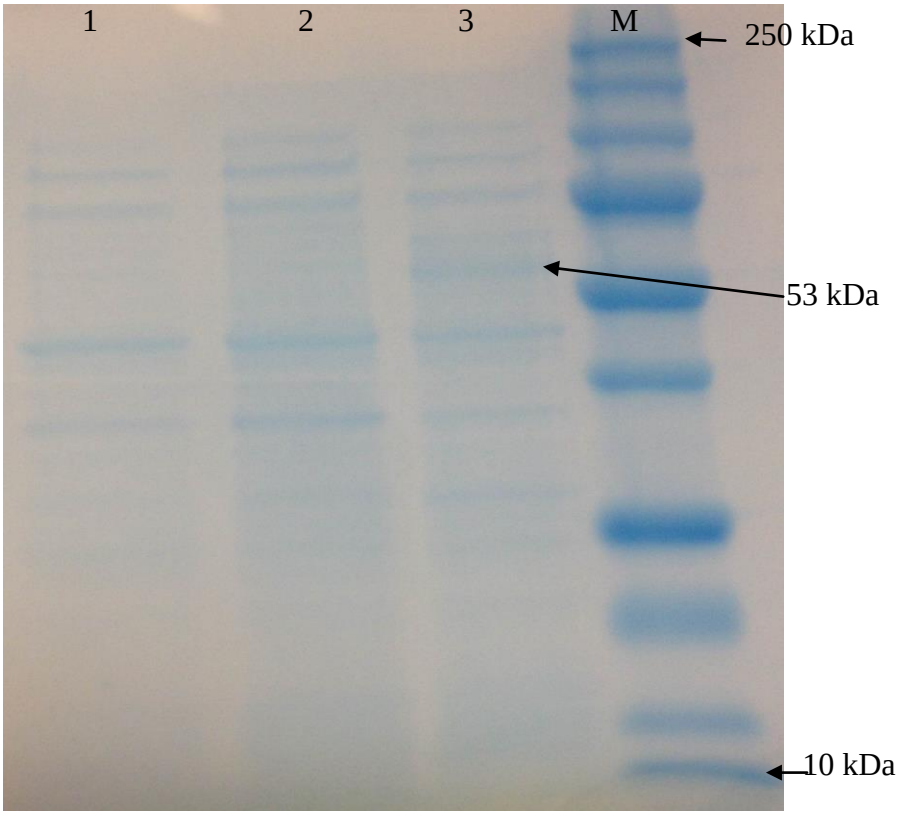


Figure 4.1 SDS-PAGE analysis of the cell lysate and soluble fraction from transformed *S. blattae* containing the recombinant plasmid with Dehly_1524 grown at 18 ºC. Expected sized (53 kDa) protein band observed with cell lysate (lane 3). Lane 2 is the uninduced transformed *S. blattae* cells whereas Lane 1 is the soluble fraction.

4.3.3 Enzyme assay results from the induced proteins

No enzyme activity was observed from the large scale protein induction (using 250-1000 mL of cell culture volume) using the transformed *S. blattae* cells. Induced protein (crude lysate) in the concentration range of 250-750 µg was used for the dehalogenase enzyme assay; however, using this range of protein did not show functional activity. For every assay performed, a fraction of the induced cell culture was analyzed by 12% SDS-PAGE. The expected sized protein was consistently observed in the SDS PAGE analysis in the insoluble fraction. 10-15%

dechlorination was consistently observed in positive controls with cell lysates from *D. lykanthroporepellens* BL-DC-9^T.

The Mac Nelly *et al.,* group also used a range of 250-750 µg of protein in the dehalogenase assay. Dehalogenase activity was detected when this range of protein was tested in the enzyme assay. Recently a functional dehalogenase was expressed using *E. coli* as the host. The purified and cofactor reconstituted vinyl chloride reductive dehalogenase at a concentration 0.2-1.0 mg/mL was shown to exhibit dehalogenase activity (Parthasarathy *et al*., 2015).

4.3.4 Enzyme assay results from the ligand binding experiments

For the enzyme assays using mixtures of induced lysates of transformed *S. blattae* and *D. lykanthroporepellens* BL-DC-9^T, approximately less than 50% enzyme activity was observed in the dehalogenase enzyme assay when the lysates of *S. blattae* and *D. lykanthroporepellens* BL-DC-9^T were mixed in the ratio 1:1 in comparison to the positive controls using only BL-DC-9^T cell lysate.

4.4 Discussion

Because the first functional expression of a reductive dehalogenase was reported in the cobamide producing *S. blattae*, it was speculated that it may work in the functional expression Dehly_1524. The tetrachloethene reductive dehalogenase (PceA) of the *Desulfitobacterium hafniense* strain Y51 and *rdhA3* of *D. hafniense* DCB-2 was functionally expressed in the *S. blattae* cells (Mac Nelly *et al*., 2014). Hence, functional expression of the 1,2-DCP reductive dehalogenase protein from strain BL-DC-9^T was attempted using *S. blattae* as the host. The gene was synthesized along with the TAT signal sequence because when the first functional

expression of a dehalogenase was successfully carried out, it included the full-length reductive dehalogenase gene (Mac Nelly *et al*., 2014).

The 53 kDa 1,2-DCP reductive dehalogenase protein (Dehly_1524) was found to be expressed consistently in *S. blattae* cells under anoxic conditions in the insoluble fraction. The (PceA) of the *D. hafniense* strain Y51 was also found to be expressed in the insoluble fraction in the same *S. blattae* host, and functional activity of this enzyme was reported in spite of the protein being expressed in the insoluble fraction (Mac Nelly *et al*., 2014). Co-expression of the PceT gene reportedly increased the solubility and functional activity of this enzyme (Mac Nelly *et al*., 2014). In all attempts that were made to express the Dehly_1524 protein in experiments reported here, the protein was expressed in the insoluble fraction and it was found to be enzymatically inactive in spite of using a range of induction times and a range of anhydrotetracycline concentrations for induction.

The insert gene *PceT* was sequenced at GenScript (Piscataway, NJ) after the subcloning in the pASK-IBA 3C vector and the sequencing results were consistent with the targeted gene sequence. However, attempts to co-express the *PceT* gene from *D. hafniense* along with the *dcpA* gene from *D. lykanthroporepellens* in the *S. blattae* cells were not successful. It remains a possibility that although the DcpA protein was being made in the *S. blattae* cells, the protein was nonfunctional because the oxidation reduction potentional was outside the limit for the functional activity of the enzyme. Also, it may be that 5,6-dimethylbenzimidazole (DMB) and hydroxocobalamine are not the cofactors required for the functional activity of the enzyme. Alternatively, it is also possible that the *de novo* cobamides synthesized in the *S. blattae* cells do not function as cofactors, and hence, it fails to produce the mature form of the enzyme. As mentioned in the previous chapter, cobamide extraction and analysis has shown that the *S.*

blattae cells could produce pseudo-B_{12} that has dimethylbenzimidazole or adeninyl as the lower ligand. The dimethylbenzimidazole or the adeninyl moiety may not be the required cofactor for the DcpA enzyme. It remains likely that the *de novo* cobamides synthesized in the *S. blattae* cells do not function as cofactors, and hence, fail to produce the mature form of the enzyme. The cobamide cofactor that may actually be needed for the functional activity of the DcpA protein still remains unknown and is possibly not produced endogenously by the *S. blattae* cells and was not added to the cell culture.

Refolding of the vinyl chloride reductive dehalogenase protein from the insoluble fraction followed by iron sulfur cofactor reconstitution in the presence of hydroxocobalamin has shown functional activity (Parthasarathy *et al*., 2015). It remains to be investigated if the iron sulfur cofactors were reconstituted into the DcpA protein during expression in *S. blattae* cells. The possibility that the iron sulfur cluster was not incorported into the enzyme cannot be ruled out. The reconstituted iron sulfur cluster could possibly mediate an electron transfer, which could possibly be a critical step during the dehaloelimination process. Hence failure to reconstitute the iron and sulfur cofactors in the DcpA protein during expression in *S. blattae* could be a plausable reason for the expression of a catalytically inactive enzyme.

In previous attempts to express a functional dehalogenase in *E. coli*, the absence of a *de novo* cobamide biosynthesis pathway was speculated to be a likely reason for the failure to produce a functional protein. Work reported herein, however, indicates that presence of a *de novo* pathway is not sufficient for the functional expression of this family of proteins. In an attempt to resolve this potential issue, experiments were carried out to test if mixing the cell lysates of *D. lykanthroporepellens* BL-DC-9^T and *S. blattae* cells could produce a functional enzyme. The cell lysates of *S. blattae* and *D. lykanthroporepellens* BL-DC-9^T were mixed in the

ratio 1:1 and used in the dehalogenase enzyme assay and the enzyme activity was compared to the postive controls using only BL-DC-9^T cell lysate. Dehalogenase enzyme activity in the mixture of induced *S. blattae* and *D. lykanthroporepellens* BL-DC-9^T, however, was approximately 50% lower than the positive controls containing cell lysates of only *D. lykanthroporepellens* BL-DC-9^T, indicating that mixing the *S. blattae* and *D. lykanthroporepellens* BL-DC-9^T did not resolve the ligand binding issue. Functional dehalogenase activity was recently observed by Parthasarathy *et al*., 2015 using a very high concentration of purified and cofactor reconstituted vinyl chloride reductase (0.2-1.0 mg/ml). However the concentration of protein in the cell lysates of *S. blattae* was less than than the concentration of purified enzyme that exhibited functional activity. Hence it remains a possibility that catalytically active enzyme was produced but the amount of protein used in the assay was not adequate to show a dehalogenase activity.

Hence it can be concluded from these experiments that while *S. blattae* can be used for the functional expression of some reductive dehalogenase proteins (i.e., those from *D. hafniense* as reported by Mac Nelly *et al*., 2014), it may not work for all proteins of this family.

4.5 Conclusions

Since functional expression of the 1,2-DCP reductive dehalogenase protein was not possible in *E. coli*, with absence of the cobamide cofactor most likely a contributing factor, functional expression of the protein was attempted in the cobamide producing *S. blattae* cells. The protein was expressed in the insoluble fraction, and functionally active protein was not obtained. Also, mixing the *S. blattae* and *D. lykanthroporepellens* BL-DC-9^T cell lysates did not produce a catalytically active enzyme. Hence, it can be concluded that although *S. blattae* has been successfully used as a host for functional expression of (PceA) of the *D. hafniense* strain

Y51 (Mac Nelly *et al*., 2014) this species should not be assumed to be a universal host for expression of other dehalogenase proteins. Heterologous expression of functional reductive dehalogenases remains a challenge.

Chapter 5. Overall conclusions and recommendations for future research

5.1 Overall conclusions

Transcriptomic studies were performed to study the pattern of gene expression in *Dehalogenimonas lykanthroporepellens* BL-DC-9^T during organohalide respiration in presence of 1,2-DCA, 1,2-DCP and 1,2,3-TCP as electron acceptors. Using the genome sequence available for this strain, specific primer pairs were designed targetting all 25 *rdhA* gene sequences present in BL-DC-9^T. For six of the *rdhA* genes having the cognate *rdhB* genes, primers were designed that spanned both genes, and this was used to check if the two genes were cotranscribed. RNA extracted from actively dechlorinating cultures of BL-DC-9^T was reverse transcribed, and expression of the *rdhA* and *rdhAB* genes was detected by polymerase chain reaction in conjunction with designed primers. From this study, it was observed that the pattern of gene expression in *D. lykanthroporepellens* BL-DC-9^T was identical during organohalide respiration with all three electron acceptors that were used in this study (i.e., 1,2-DCA, 1,2-DCP, and 1,2,3-TCP). Transcripts from nineteen *rdhA* genes were consistently detected from the RT-PCR experiments irrespective of the electron acceptor used. This included thirteen full-length *rdhA* genes, 6 truncated *rdhA* genes, and 5 *rdhA* genes having the cognate *rdhB* gene. The five *rdhB* genes were also found to be co-expressed with their corresponding *rdhA* genes. The six *rdhA* genes for which transcripts were not detected in any of the RT-PCR experiment appear unlikely to play a role in the organohalide respiration process using the substrates tested here.

The three transcripts lacking the TAT signal sequences and detected in the gene expression study Dehly_0069, Dehly_1523 and Dehly_1582 are interesting candidates for functional heterologous productions. They appear to be expressed under all the tested conditions and they lack the TAT signal sequence and the membrane anchoring *rdhB* gene indicating they

are not membrane proteins. Hence they can possibly be expressed in the soluble fraction both in *E. coli* and *S. blattae*. The six transcripts that were not detected in the gene expression study included 4 full-length *rdhA* genes (Dehly_0068, Dehly_0283, Dehly_1355 and Dehly_1540) and two truncated *rdhA* genes (Dehly_0075 and Dehly_1533/Dehly_1534). Both Dehly_0075 and Dehly_1533/1534 lack the TAT signal sequence and hence both possibly code for truncated cytoplasmic proteins having no role in the organohalide respiration process. That some full-length *rdhA* genes were not transcribed likely reflects that they are under control of transcription factors that respond to ligands other than the compounds used in this study.

Transcripts from the four regulatory genes that are associated with the *rdhA* genes were detected in this study, indicating their involvement in regulation of gene expression during the biotransformation process. The majority of the *rdhA* genes in *Dehalococcoides mccartyi* are associated with either two component system (TCS) regulators or MarR regulators (Kube *et al*., 2005; Seshadri *et al*., 2005). However, the genome analysis shows the presence of only four regulatory genes adjacent to the *rdhA* genes in *D. lykanthroporepellens* BL-DC-9^{T}. The regulatory genes may possibly regulate other *rdhA* genes and not only the adjacent *rdhA* genes. Hence, it can be inferred that the two component systems transcription regulator Dehly_1149 located adjacent to the *rdhA* Dehly_1148 may be involved in regulation of other *rdhA* genes such as Dehly_1152 and not only the adjacent *rdhA* genes. The regulation of the *rdhA* genes is not compound specific and rather promiscuous, an observation that must derive from the ligand-specificity of associated regulators. The genome sequence shows the presence of a MarR orthologue Dehly_0120 located immediately upstream and in opposite orientation of RdhA Dehly_0121. A pseudo-palindromic sequence (TCGATGTAATATTACATGTA) with two half sites for MarR binding were found to be located immediately upstream of the Dehly_0121 gene,

strongly suggesting that this site serves as the MarR binding site. The genome sequence of BL-DC-9 was analyzed using pattern locator software (Mrazek and Xie, 2006) for MarR binding site and the analysis showed that similar palindromic sequences were not present upstream of any of the other *rdhA* genes. Hence it is concluded that the MarR orthologue likely regulates the expression of Dehly_0121 and not other *rdhA* genes.

The remaining transcription factors that are encoded adjacent to *rdhA* genes are part of two-component systems. Sensor kinases would be expected to respond to specific signals (the presence of halogenated compounds) by transferring a phosphoryl group to the response regulator, which in turn controls gene expression. These response regulators also belong to the helix-turn-helix family of DNA binding proteins and would be expected to recognize palindromic sequences. Searching the genome with the palindromic sequence TACAATTATTGTA found in the Dehly_1152 (*rdhA*) promoter uncovered similar sites in the Dehly_1148 and Dehly_0156 promoters. This raises the possibility that these three genes are controlled by the same regulator, perhaps the two-component regulator encoded by Dehly_1149-1150.

Although it was not possible to uniquely identify the enzyme-coding genes from this study, this study does provide insights into the gene expression pattern during the dechlorination process in this interesting organism. The primers developed in this study can be used in future studies for quantifying the transcripts using a qRT-PCR approach. Also the primers can be used as probes for checking the presence of these genes from environmental samples and have potential utility in assessing *rdh* gene expression at contaminated sites.

Once the pattern of gene expression was identified, the next study was done to heterologously express a catalytically active dehalogenase enzyme. One of the *rdhA* genes whose

transcripts were detected consistently during the dechlorination of 1,2-DCA, 1,2-DCP and 1,2,3-TCP has been characterized as the gene responsible for the biotransformation if 1,2-DCP. This gene, with the locus tag Dehly_1524, was chemically synthesized and codon optimized for expression in *E. coli*. It was subcloned in pET-16b and transformed into *E. coli* BL21(DE3)pLysS cells for expression. In spite of many attempts, however, the protein was found to be functionally inactive and the majority of the protein was expressed in the insoluble fraction. Refolding the protein from the insoluble fraction in the presence of the vitamin B_{12} did not produce a functional protein. The inability to achieve functional expression of the protein in *E. coli* is likely due to the absence of a *de novo* cobamide biosynthesis pathway in the *E. coli* cells.

Because the first successful expression of a reductive dehalogenase protein was reported using *S. blattae* and since *E. coli* was found to be unsuitable due to the absence of a cobamide biosynthetic pathway, expression was tried using *S. blattae* as an alternate host (Mac Nelly *et al*., 2014). The presence of a *de novo* cobamide biosynthesis pathway in *S. blattae* makes it a promising host for testing the functional expression of dehalogenases. Using the same approach by the Mac Nelly *et al*. (2014) group, the Dehly_1524 protein was expressed under anoxic conditions in *S. blattae* cells. The functional protein could not be obtained, however, and despite all attempts, the protein was mostly expressed in the insoluble fraction. Attempts to co-express the PceT protein along with the Dehly_1524 protein were not successful with *S. blattae* cells. This indicates that *S. blattae* may not function as a genetic host for the expression of all dehalogenase enzymes, particularly those from the genus *Dehalogenimonas*. It could be that the cobamides produced by *S. blattae* differ in the lower axial ligand from the cobamides used by *D. lykanthroporepellens* BL-DC-9^T, resulting in a catalytically inactive enzyme. Since a mixture of

S. blattae and *D. lykanthroporepellens* lysates did not result in restoration of functional activity, it is also likely that the cobamide cofactor must be present during protein folding.

5.2 Recommendations for future research

To obtain further insights into the *rdhA* gene expression, the transcript levels of the nineteen detected *rdhA* genes could be quantified using quantitative RT-PCR or using microarrays. Because the heterologous expression of a functional dehalogenase was not successful in the experiments described here, a wide scope exists for method development for successfully producing functional dehalogenases from *D. lykanthroporepellens* BL-DC-9^T. Attempts to try different cobamides as cofactors represents one option, and trying to express the protein in an even more reducing environment remains another possibility. It remains to be investigated whether reconstituting the Fe/S clusters in the expressed proteins might result in the production of functionally active proteins. Once a method is developed for functional expression of reductive dehalogenases, the enzyme's substrate specificity can be studied by testing dechlorination activity on different halogenated electron acceptors. Enzyme kinetics of a functional enzyme would obviously be worth study, and additional experiments are needed for the purification of the protein. Determining the X-ray structure of the protein for prediction of the active sites of the protein would be a logical step.

Additional research is required for characterization of several other *rdhA* genes that were detected in the RT-PCR experiments. In total, five *rdhA-rdhB* operons were detected in the gene expression study. It is needed to check if the other *rdhA* genes such as Dehly_0156, Dehly_0275, Dehly_1152 and Dehly_1054 can be functionally expressed heterologously. It may be possible that the other detected dehalogenase genes are actually enzyme coding genes catalyzing the

transformation of 1,2-DCA or 1,2,3-TCP. Functional expression of the other dehalogenase genes could provide further insights into the dechlorination process in this organism.

As a means of providing a potential roadmap for future efforts to overcome protein misfolding in hetologous expression of reductive dehalogenases, the following sub-sections review some alternative approaches that may prove successful in resolving DcpA protein aggregation.

5.3 A review of different approaches that can be potentially be used to resolve the DcpA protein aggregation

Protein misfolding or aggregation is one of the most commonly observed obstacles in structural genomics and proteomics research (Yokoyama, 2003). Many interesting families of proteins such as membrane proteins, kinases and phosphatases are extremely onerous in getting expressed in the soluble fraction of *E. coli* cells (Esponsito and Chatterjee 2006). To obtain such recombinant proteins in their soluble (i.e., non-aggregated) form, a wide variety of techniques have been employed. As described in the following sub-sections, these approaches can be broadly categorized as (1) changing the growth conditions of the transformed cells and altering the rate of protein synthesis, (2) changing the expression host, (3) use of solubility fusion tags. In addition statistical methodologies to prevent protein aggregation and some experimental approaches for protein refolding from the insoluble fraction are mentioned in section 5.3.4 and 5.3.5.

5.3.1 Changing the growth conditions of the transformed cells and altering rate of protein synthesis

One of the primary factors that can improve the solubility of recalcitrant proteins is by altering the media or culture composition. The concentration of peptone and yeast extract in the

protein induction media are known to have an affect on the protein solubility (Larentis *et al*., 2011). Several compounds are reported that when added to the culture medium result in proper protein folding and increase the solubility. These include sorbitols, thiols, ethanol and disulfides (Georgiou *et al*., 1996, Chopra *et al*., 1994).

The next factor that often inhibits protein aggregation in expression hosts is altering the rate of protein synthesis in the expression hosts. Induction of the protein expression in the late log phase and in some instances in the stationary phases have led to the production of the protein in the soluble fraction (Galloway *et al*., 2003, Ou *et al*., 2004). Mammilian RNA binding protein has been successfully expressed in high quantity in the soluble fraction in *E. coli* cells by using lower concentration of IPTG at the late log phase (Galloway *et al*., 2003). The induction temperature and the length of the induction time can affect both the protein expression and the protein solubility (Papaneophytou and kontopidis, 2014). Induction of the proteins at temperatures below the optimal temperatures for cell growth has been reported for expression of soluble proteins in bacterial hosts as lower temperatures slow down the cell growth and protein expression (Schein *et al*., 1988). At this point, the optimum temperature and the length of the induction time for a particular protein can only be determined by trial and error.

Expression of the target proteins without using an external inducing agent can also be a favourable option for the expression of tricky proteins in the soluble fraction. In this approach, generally an autoinduction media is used to express target proteins placed under the control of a *lacUV* promoter and this system was used for the expression of selenomethionine labeled fusion proteins in *E. coli* cells (Sreenath *et al*., 2005). Both glucose and lactose are supplemented to the autoinduction media, and leaky expression of the target protein is prevented by catabolite repression. When glucose from the autoinduction media is used up, the lactose

metabolite allolactose is produced that acts as the inducer for the *lacUV* promoter and leads to the expression of target proteins (Sreenath *et al*., 2005).

The pET16b vector was used for the expression of the DcpA protein in this study. The pET16b uses an IPTG-inducible, bacteriophage T7 promoter that has been widely used for expression of a large number of proteins effectively (Tabor *et al*., 1985). It was observed, however, that the majority of the DcpA protein was expressed as inclusion bodies. Hence, this expression may not be suited for preventing protein aggregation of the DcpA protein. Other than the above mentioned T7 expression system, cloning and expression using the tightly controlled araC promoter (Guzman *et al*., 1995) that is inducible by arabinose remains a favorable option for producing protein in the soluble fraction. A switch to a weaker expression system, such as a *lac* promoter may lead to the slower expression of the desired protein and hence may be useful for rectifying the protein aggregation that results from faster protein expression (Labendiker and Danieli, 2014).

Slower expression of proteins using a cold shock promoter like *cspA* has also been reported to solve the protein aggregation problem (Vasina *et al*., 1996). Using lower temperatures for the expression of difficult proteins have often circumvented the inclusion body formation problem (Schein, 1989). Still a significant percentage of proteins are expressed as inclusion bodies when expressed at sub-optimum temperatures (15-30 ^{o}C) in *E. coli* systems (Song *et al*., 2012).

A number of protein solubilizing agents, osmolytes and key additives have been characterized as agents that increase solubility of misfolded proteins (Bondos *et al*., 2003, Leibly *et al*., 2012, Jain and Roy, 2009). Some of the media additives such as CTAB, K_2PO_4, proline, glycine, potassium citrate, xylitol, $CuCl_2$ are among the important protein solubilization agents

(Leibly *et al*., 2012). L-arginine has been widely used as a protein solubilization additive since it has been reported to subdue protein aggregation (Arakawa *et al*., 2007).

5.3.2 Changing the expression host

Another factor that can affect the solubility of the recombinant protein in the choice of the expression host. The *E .coli* BL21(DE3) system is usually the standard genetic host system for the production of recombinant proteins (Shaw *et al*., 1967). Membrane-bound proteins that are difficult to express in the soluble fraction of BL21(DE3) cells have been expressed without inclusion body formation in mutant *E. coli* strains C41(DE3) and C43(DE3) (Miroux *et al*., 1996). Preference of using these mutant cells as the expression host over the standard BL21 cells can be a promising choice for expression of the membrane-bound proteins. The recently genetically engineered *E. coli* mutant Lemo 21(DE3) strain has also been reported as a preferable choice for membrane protein solubilization (Hjelm *et al*., 2013). The *Ralstonia eutropha* expression system has been reported as a higher ranking expression system in comparison to the standard *E. coli* system as an expression in this system minimizes inclusion body formation (Barnard *et al*., 2004). Several strains of the genus *Bacillus* have been reported to act as better host for the expression of recombinant proteins (Demain and Vaishnav, 2008). *Bacillus megaterium, B. subtilis* and *B. licheniformis* are some of the species of the genus *Bacillus* that are routinely used for heterologous protein expression (Demain and Vaishnav, 2008). Hence, it can be concluded that apart from the standard *E. coli* system other bacterial expression systems may be used for expression of recombinant proteins that are found to be misfolded in the *E. coli* system.

Use of extremophilic *E. coli* strain such as *E. coli* ArcticExpres as expression hosts has resulted in successful expression of soluble proteins at temperatures below 10 ºC (Ferrer *et al*.,

2003). Protein aggregation and misfolding was prevented when the expression host *E. coli* Rosettagami2 (DE3) was grown at a temperature below 10 ºC (Song *et al*., 2012). This specific study exhibited that protein misfolding for difficult proteins such as *Cryptopygus antarticus* mannase can be prevented using low induction temperatures (6-10 ºC). This recent study is a significant refinement of the study of recombinant proteins that are expressed as aggregates at optimal temperatures for *E. coli* growth. Hence using *E. coli* Rosettagami2 (DE3) strains for the expression of the DcpA protein at a sub-optimal temperature of (6-10 ºC) seems to be a plausible option to express the protein in the soluble fraction.

5.3.3 Use of solubility fusion tags

One of the key factors that have lead to a significant improvement if the field of protein expression and solubilization is the development and use of solubility tags to reduce protein aggregation. GST (Glutathione S-transferases) and maltose binding affinity tags are the most frequently used tags for solubility enhancement of proteins that have low solubility and tend to form inclusion bodies (Yokoyama 2003, Smith *et al*., 1988). The GST/Bombyx mori nucleopolyhederovirus DNA polymerase fusion protein has resulted in higher quality of heterologous expression of this protein in *E. coli.* The GST/Bombyx mori nucleopolyhederovirus DNA polymerase fusion protein was further purified using affinity chromatography (Song *et al*., 2014). Similarly elastin like polypeptides that were mostly expressed as aggregates when expressed in *E. coli* cells were found to be expressed in the soluble fraction when fused to the maltose binding protein (Batallie *et al*., 2015). The use of the GST tag however often leads to the expression of truncated proteins and thereby fails to resolve the protein aggregation completely (Labendiker and Danieli, 2014). Maltose binding tags function not only in protein solubilization but also act as protein stabilizers (Labendiker and Danieli, 2014). However, a severe

disadvantage of this solubility tag is that it has a large size and hence its use is limited only to the solubilization of smaller proteins (Labendiker and Danieli, 2014). The *E. coli* expression system is ineffective at proper folding and solubilization of large sized fusion proteins (Labendiker and Danieli, 2014). Hence, larger sized proteins are generally misfolded when fused to a maltose binding tag and expressed in an *E. coli* system. The NusA (N utilization substance A) tag has also been reported to be an effective fusion protein that prevents protein aggregation. Catalytically active human interleukin was expressed as a fusion protein in the soluble fraction usig a fusion NusA tag and was purified using metal affinity chromatography (Davis *et al*., 1999). The use of the NusA tag as a protein solubilizer in *E. coli* also suffers from a similar drawback like the maltose binding tags. In spite of having higher solubility than maltose binding tags, a large size of 55 kDa prevents its effective usage in the expression of larger proteins in bacterial systems (Nallamsetty and Waugh, 2006). The other major drawback of these solubility tags is that they are not suitable for the NMR spectroscopy of tagged proteins and cleaving out these tags before NMR spectroscopy often tends to decrease the protein solubility (Yokoyama, 2003).

A development in the field of NMR spectroscopy for proteins with low solubility has been the development of the non-cleavable N-terminal G fusion protein (B1 domain) (Zhou *et al*., 2001). This tag has been developed as a non-cleavable solubility tag that does not interfere with the functionality of the protein and is hence useful for determination of protein structure (Zhou *et al*., 2001). It has been reported that the GFP (Green fluorescent protein) can be used as the N-terminal fusion tag to rectify protein misfolding and enhance the solubility of proteins (Pedelacq *et al*., 2002). Using GFP as the N-terminal fusion tag led to the expression of a hyperthermophilic protein successfully in *E. coli* (Pedelacq *et al*., 2002). The fluorescence

signals from GFP-tagged proteins in the *E. coli* host cells is a signal to confirm the correct refolding of misfolded proteins and the GFP tag serves as the folding reporter (Waldo et al., 1999). In many instances study of the interactions of low soluble proteins with other proteins, nucleic acids and ligands are critical for enhancing the solubility of the proteins. Co-expression of associated chaperones and refolding proteins can result in the expression of the protein in the soluble fraction. Co-expression of bacterial periplasmic chaperones such as thiol disulfide oxidoreductases and peptidyl-prolyl cis trans isomerases has been effective in production of soluble recombinant proteins in bacterial systems (Schlapschy *et al*., 2011). The first instance of heterologous production of a soluble reductive dehalogenase was the *pceA* gene from *Dehalobacter restrictus* (Sjuts *et al*., 2012). The soluble protein was expressed in an *E. coli* host system by fusing the PceA with a strep tag as a solubility tag (Trigger factor) and by expressing the protein at cold temperatures. The fusion protein was purified using a StrepTactin Sepharose resin. The tetrachloethene reductive dehalogenase (PceA) of the *Desulfitobacterium hafniense* strain Y51 was functionally expressed in the *S. blattae* cells. This specific study exhibited that the production of the soluble enzyme increased when the trigger factor chaperone PceT located downstream was co expressed along with the *pceA* gene (Mac Nelly *et al*., 2014).

In addition to the above-mentioned affinity tags there are numerous other tags that have been successfully used for increasing the solubility of misfolded proteins. These include calmodulin binding peptide (Wyborski *et al*., 1999), and thioredoxin (trx) (LaVallie *et al*., 1993). Calmodulin binding proteins have binding affinities for calmodulin and can be purified using calmodulin affinity chromatography. Heterologous expression and purification of the c-Jun-N terminal kinase protein has been reported in *E. coli* using an N terminal calmodulin binding protein (Wyborski *et al*., 1999). The thioredoxin fusion tag has led to the functional production

of eukaryotic cytokines in the soluble fraction in *E. coli* cells (LaVallie *et al*., 1993).The solubility fusion tags are frequently used along with a simple histidine tag that functions as affinity tags to permit the purification of soluble proteins (Porath, 1992). Other frequently used N-terminal fusion tags for enhancing protein solubility includes fusion of the SUMO protein (Butt *et al*., 2005), HaloTag7 (Ohana *et al*., 2009) and ubiquitin (Catanzariti *et al*., 2004). N terminally fused SUMO (Small ubiquitin related modifier) tags have been used for the functional expression of green fluorescent proteins (Malakhov *et al*., 2004). The major edge of the SUMO tag is that it is easily cleavable from the fusion protein using SUMO proteases and has a higher efficiency for increasing solubility of larger proteins due to its small size (Labendiker and Danieli, 2014). The HaloTag7 that is derived from a nonfunctional haloalkane dehalogenase from *Rhodococcus* species (Janssen, 2004) is one of the most effective tag in protein solubilization. Experimental testing of the HaloTag7 as a fusion protein to 23 human proteins have produced soluble heterologous production of 74% of these proteins in *E. coli* (Ohana *et al*., 2009).

Inclusion body formation is also a major obstacle in the functional expression of eukaryotic proteins in *E. coli* (Chatterjee and Esponsito, 2006). Successful expression of human proteins such as IL 13 and folliculin in *E. coli* using tags derived from bacteriophage T7 protein kinase and small *E. coli* chaperone skp have been described (Chatterjee and Esponsito, 2006). The remarkable advantage of these two tags over other previously reported tags was that the expressed proteins remained in the soluble fraction even after removal of the tags through protease cleavage (Chatterjee and Esponsito, 2006). Hence, using bacteriophage T7 protein kinase and small *E. coli* chaperone skp as fusion tags while expression in *E. coli* remains an option for the expression of truculent proteins that otherwise tend to aggregate during expression.

In the absence of experimental testing, identification of a specific tag that may function as a solubilizer for an uncharacterized protein is difficult. Hence, tag switching is routinely performed in order to identify the tag(s) that may increase the protein solubility. In spite of the success of solubility tags for increasing protein solubility, major challenges remain with the use of solubility tags (Labendiker and Danieli, 2014). These include the high costs associated with the removal of the solubility tags post protein expression and loss in protein solubility and functional activity post removal of the solubility tags (Labendiker and Danieli, 2014).

It can be seen from this review that there are numerous modifications (e.g., changing media compositions, switching between solubility tags, using different promoters and different host system and induction for shorter durations and at lower temperatue) that could be made to the expression system to prevent potential protein misfolding. An approach that combines many of these modifications together has also been reported to result in enhanced protein solubility (Labendiker and Danieli, 2014). In this approach, a preliminary analysis of the target proteins are undertaken using bioinformatics tools to identify the protein characteristics (such as hydrophobic, membrane-bound, intrinsically disordered, associated with chaperones etc). The analyzed proteins are then fused to N-terminal solubility enhancement tags depending on the size of the protein (SUMO tags for larger sized proteins and maltose binding tags for smaller sized proteins). This approach tests at least four different *E. coli* host strains as the protein expression host. Apart from the regular BL21 (DE3) and HMA174 (DE3) hosts strain C41 (DE3) and C43 (DE3) are all transformed with the solubility tag-target fusion construct. The transformed recombinant cells are subsequently innoculated in an autoinduction media for 16 hours and induced at different temperatures. After the induction, the induced cells are lysed and expression of the target proteins are verified by SDS-PAGE and Western blot analysis. After determination

of the best possible host-construct combinations, additional factors such as induction temperature and induction durations can be determined by trial and error. This strategy has proven effective to minimize protein aggregation for a lot of difficult proteins (Labendiker and Danieli, 2014).

5.3.4 Statistical methodologies to prevent protein aggregation during heterologous expression

Because trial-and-error testing to determine which combinations of many factors may prevent protein aggregation remains a very laborious job, different statistical approaches have been recently devised to provide guidance in combining different factors that affect protein solubility (Singh *et al*., 2009). Statistical design of the experiments uses factorial experimental models to reduce the total number of experiments (Chambers *et al*., 2009).

An effective statistical approache to minimize protein aggregation is to use a full factorial design that uses all possible experimental factors affecting protein solubility. If *k* is the number of factors for a particular experiment and each factor has an upper and lower level, then 2^k is a possible 2-level full factorial design. Using this full factorial model, multiple simulation runs can be made for each of the factors, and results for the combined effect of all the factors can be obtained (Chambers *et al*., 2009). A partial factorial design is used in the case of a large number of experimental factors. A partial factorial design uses a large number of factors, however with reduced number of variables and only a fraction of the total experimental conditions are analyzed. The statistical approach of full and partial factorial design has been successfully used to identify several potential factors affecting protein solubility such as inducer concentrations, induction time, cell line and media and the effects of using protein solubilizers (Benoit *et al*., 2007). As innumerable variable factors such as expression host, solubility fusion tag, induction temperatures, culture medium can have major and minor effects in protein aggregation. It is

actually almost impossible to identify the effects of each of the variables; hence, only the factors that have major effects should be taken into consideration. The total and partial factorial design approach helps in the identification of the major effects of the variable factors (Abergel *et al*., 2003). The data collected from the total and partial factorial design approach can be analyzed using available statistical methods such as ANOVA (univariate analysis of variance), MANOVA (multivariant analysis of varient) and linear coefficient analysis (Papaneophytou and Kontopidis, 2014).

In addition to the factorial design models, increase in protein solubility has also been achieved using the response surface methodology (Tabandeh *et al*., 2008). Response surface methodology in combination with factorial design model was used for the expression of multiple eukaryotic proteins in bacterial expression systems (Papaneophytou et al., 2013, Papaneophytou et al., 2014). In these studies, a statistical analysis identified BL21 (DE3) pLysS as the most suitable expression host. Subsequently the optimum optical density for induction, inducer concentration, induction temperatures and duration of induction to minimize protein aggregation was achieved by response surface methodology. In a separate study, the effect of several factors on protein solubility of the recombinant zinc-metalloprotease was investigated using fractional factorial study in combination with response surface methodology (Beigi *et al*., 2012).The fractional factorial study was used to investigate the effect of the variables (induction temperature and induction duration, media composition, the optical density of the culture at the time of induction and Ca^{2+} concentration and inducer concentration. After the initial analysis, response surface methodology was used to examine the effects of the three variables (Ca^{2+} concentration, inducer concentration and optical density of the culture at the time of induction.

5.3.5 Potential refolding experiments that may be attempted for prevention of the DcpA protein aggregation

For proteins that tend to aggregate and misfold under all tested expression conditions, refolding of the protein from the insoluble fractions can be used to obtain the active form of the enzyme. In this method, the aggregated proteins are dissolved using denaturants such as guanidium hydrochloride or urea and then refolded using dilution, dialysis or column chromatography (Jobby et al., 2006). The D2 domain of *Yersinia crystalline* was overexpressed using *E. coli* as the host, and majority of the protein was expressed in the insoluble fraction. Active form of the domain was purified using refolding by column chromatography, and a method was developed to separate nucleic acids from co elution with the protein. The method developed led a rapid purification of the enzyme with a higher protein yield in comparison to that purified from the soluble fraction (Jobby *et al*., 2006).

An alternative method of protein refolding to that of the traditional methods of dilution or dialysis is the matrix assisted refolding. A eukaryotic heat shock protein Hsp26 from *Saccharomyces cerevisiae* was expressed in *E. coli* BL21 (DE3) cells in the form of aggregates and then refolded to its active form using a matrix assisted refolding. The protein was refolded in an immobilized state (Franzmann 2006). Refolding of the protein from the insoluble fraction was also used in the biochemical characterization of the esterase EM2L8 from metagenomic library of uncultured microorganisms from deep sea sediments. The protein was expressed in the insoluble fraction in the *E. coli* hosts at an induction temperature of 37 ºC and in the soluble fraction at an induction temperature of 18 ºC. The insoluble fraction also exhibited enzymatic activity on gel during SDS-PAGE electrophoresis, suggesting that the protein was refolded during the electrophoresis. The insoluble fraction showed activity after denaturation with a

detergent followed by dilution that confirmed that the protein was refolded to the active form from the misfolded state (Park *et al*., 2007).

Instead of using denaturants to solubilize the inclusion bodies, non denaturating agents can also be used for expression of protein in the soluble fraction. Non denaturing agents are particularly useful for the expression of proteins that tend to get expressed partly in the soluble and partly in the insoluble fraction. Partial expression into the insoluble fraction is due to slight misfolding in comparison to the native structure of the proteins and many protein families such as kinases, tumor necrosis factors, and interferon-gamma have been expressed partially in insoluble fraction (Tsumoto *et al*., 2010). Use of arginine as a non denaturing solubilization agent has prevent protein aggregation of a lot of proteins (Tsumoto *et al*., 2010). Inclusion body formation is also caused by the formation of the native like secondary protein structures in the expression host (Singh *et al*., 2005). In some cases, it has been shown that using a mild denaturing agent works better than stronger denaturants such as guanidium hydrochloride. A specific example of using a milder denaturing agent for preparation offunctionally active protein was seen in the case of the human growth hormone (Singh *et al*., 2005).

The most recent success of protein refolding dehalogenases was observed in the functional expression of a vinyl chloride reductive dehalogenase in *E. coli* (Parthasarathy *et al*., 2015). The vinyl chloride reductive dehalogenase gene was fused to an N-terminal maltose binding tag along with 6xhistidine tag and was expressed mostly in the insoluble fraction in *E. coli* BL21 (DE3) cells. The expressed protein was solubilized in 8M urea and purified using a Ni-IDA column. The purified proteins were then reduced by dithiothreitol and cofactor-reconstituted and eluted with refolding buffer in the presence of cobalamin. The cofactor-reconstituted and refolded protein showed functional dehalogenase activity. This appears to be a

promising approach for the functional expression of DcpA reductive dehalogenase and other putative reductive dehalogenases from *D. lykanthroporepellens.*

References

Abergel C, Coutard B, Byrne D,Chenivesse S, Claude JB, Deregnaucourt C, Fricaux T, Gianesini-Boutreux C, Jeudy S, Lebrun R, Maza C, Notredame C, Poirot O, Suhre K, Varagnol M & Claverie JM (2003) Structural genomics of highlyconserved microbial genes of unknown function in search of new antibacterial targets. *J. Struct. Funct. Genomics* **4**:141–157.

Arakawa T, Ejima D, Tsumoto K, Obeyama N, Tanaka Y, Kita Y & Timasheff SN (2007) Suppression of protein interactions by arginine: a proposed mechanism of the arginine effects. *Biophys. Chem.* **127**: 1–8.

Barnard GC, Henderson GE, Srinivasan & Gerngross TU (2004) High level recombinant protein expression in *Ralstonia eutropha* using T7 RNA polymerase based amplification. *Protein Expr Purif* **38**: 264–71.

Barras F, Loiseau L & Py B (2005) How *Escherichia coli* and *Saccharomyces cerevisiae* build Fe/S proteins. *Adv Microb Physiol* **50**: 41-101.

Batallie L, Dieryck W, Hocquellet A, Cabanne C, Bathany K, Lecommandoux S, Garbey S & Garanger E (2105) Expression and purification of short elastin-like polypeptides with maltose binding protein as solubility tag. *Protein Expr Purif* **110**: 165-71.

Banerjee R & Ragsdale SW (2003) The many faces of Vitamin B_{12}: catalysis by cobalamin dependent enzymes. *Annu Rev Biochem* **72**: 209-247.

Brown KL & Hakimi JM (1984) Heteronuclear NMR studies of Cobalamins. *J Am Chem Soc* **106**: 7894-7899.

Bruschi M & Guerlesquin F (1988) Structure function and evolution of bacterial ferredoxins. *FEMS Microbiol Rev* **4**: 155-175.

Beigi L, Karbalaei-Heidari H.R., & Kharrati-Kopaei M (2012) Optimization of an extracellular zinc-metalloprotease (SVP2) expression in *Escherichia coli* BL21(DE3) using response surface methodology. *Protein Expr Purif* **84**: 161–166.

Berks BC, Sargent F & Palmer T **(**2000) The Tat protein export pathway. *Mol Microbiol*.**35:** 260–274.

Benoit I, Coutard B, R Oubelaid R, Asther M & Bignon C (2007) Expression in *Escherichia coli*, refolding and crystallization of *Aspergillus niger* feruloyl esterase A using a serial factorial approach. *Protein Expr Purif* **55**: 166–174.

Bommer M, Kunze C, Fesseler J, Schubert T, Diekert G and Dobbek H (2014) Structural basis of organohalide respiration. *Science* **6208**: 455-8.

Boyer A, Pagé-Bélanger R, Saucier M, Villemur R, Lépine F, Juteau P, & Beaudet R (2003) Purification, cloning and sequencing of an enzyme mediating the reductive dechlorination of 2,4,6-trichlorophenol from *Desulfitobacterium frappieri* PCP-1. *Biochem J* **373**: 297–303.

Bondos SE & Bicknell A (2003) Detection and prevention of protein aggregation before, during, and after purification.*Anal Biochem* **316**: 223–231.

Butt TR, Edavettal SC, Hall JP & Mattern MR (2005). SUMO fusion technology for difficult to express proteins. *Protein & Expr Purif* **1**: 1-9.

Catanzariti AM, Soboleva TA, Jans DA, Board PG & Baker RT (2004) An efficient system for high level expression and easy purification of authentic recombinant proteins. *Protein Sci* **13**:1331-9.

Chopra AK, Brasier AR, Das M, Xu XJ & Peterson JW (1994) Improved synthesis of *Salmonella typhimurium* enterotoxin using gene fusion expression system. *Gene* **144**: 81-85

Chambers SP &. Swalley SE (2009) Designing experiments for high-throughput protein expression. *Methods Mol Biol* **498**: 19–29.

Cupples A M, Spormann AM, & McCarty PL **(**2003) Growth of a *Dehalococcoides*-like microorganism on vinyl chloride and *cis*-dichloroethene as electron acceptors as determined by competitive PCR. *Appl Environ Microbiol* **69:** 953–959.

Chomczynski P & Sacchi N (2006) The single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction: twenty-something years on. *Nat Protoc* **1**: 581-585.

Davis GD, Elisee C, Newham DM & Harrison RG (1999) New fusion protein systems designed to give soluble expression in *E. coli*. *Biotechnol Bioengg* **65**: 382-388.

Demain AL & Vaishnav P (2009) Production of recombinant proteins by microbes and higher organisms. *Biotechnology Advances* **27**:297–306.

De Wildeman S, Diekert G, Van Langenhove H, Verstraete W (2003) Steroselective microbial dehalorespiration with vicinal dechlorinated alkanes. *Appl Environ Microbiol* **69**: 5643-5647.

De Wildeman S, Verstraete W (2003) The quest for microbial reductive dechlorination of C_2 to C_4 chloroalkanes is warranted. *Appl Microbiol Biotech* **61**: 94-102.

Esposito D & Chatterjee DK (2006) Enhanced soluble protein expression using two new fusion tags. *Protein Expr Purif* **46**: 122-129.

Ferrer M, Chernikova TN, Yakimov MM, Golyshin PN & Timmis KN (2003) Chaperonins govern growth of *Escherichia coli* at low temperatures, *Nat Biotechnol* **21**: 1266–1267.

Fletcher KE, Ritalahti KM, Pennell KD, Takamizawa K, and Löffler FE (2008) Resolution of culture *Clostridum bifermentans* DPH-1 into two populations, a *Clostridum* sp and tetrachloroethene-dechlorinating *Desulfitobacterium hafniense* strain JH1. *Appl Environ Microbiol* **74**: 6141-6143.

Franzmann TM (2006) Matrix assisted refolding of oligomeric small heat-shock protein Hsp26. *Int J Biol Micromol* **39:** 104-10.

Fung JM, Morris RM, Adrian L & Zinder SH (2007) Expression of reductive dehalogenase genes in *Dehalococcoides ethenogenes* strain 195 growing on tetrachloroethene, trichloroethene, or 2,3-dichlorophenol. *Appl Environ Microbiol* **73**: 4439-4445.

Futagami T, Goto M & Furukawa K (2008) Biochemical and genetic basis of dehalorespiration. *Chem Record* **8**: 1-12.

Galloway CA, Sowden MP & Smith HC (2003) Increasing the yield of soluble recombinant protein expressed in *E. coli* by induction during late log phase.*Biotechniques* **34**: 524–526, 528, 530.

Gabor K, Verissimo CS, Cyran BC, Ter Horst P, Meijer NP, Smidt H, de Vos WM & van der Oost J (2006) Characterization of CprK1, a CRP/FNR-type transcriptional regulator of halorespiration from *Desulfitobacterium hafniense*. *J Bacteriol* **188**: 2604-2613.

Galperin MY (2006) Structural classification of bacterial response regulators: Diversity of output domains and domain and domain combinations. *J Bacteriol* **188**: 4169-4182.

Georgiou G & Valax P (1996) Expression of correctly folded proteins in *E. coli*. *Curr Opin Biotech* **7**: 190-197

Glod G, Angst W, Holliger C & Schwarzenbach P (1997) Corrinoid-Mediated Reduction of tetrachloroethene,trichloroethene, and trichlorofluoroethene in homogenous aqueous solution:reaction kinetics and reaction mechanisms. *Environ Sci Technol* **31**: 253-260.

Gorst CM, Yeh YH, Teng Q, Calzolai L, Zhou ZH, Adams MW& La Mar GN (1995) 1H NMR investigation of the paramagnetic cluster environment in *Pyrococcus furiosus* three-iron ferredoxin: sequence-specific assignment of ligated cysteines independent of tertiary structure**.** *Biochemistry* **2**: 600-10.

Guzman LM, Belin D, Carson MJ & Beckwith J (1995) Tight regulation,modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* **177**: 4121–4130.

Hjelm A, Schlegel S, Baumgarten T, Klepsch M, Wickstrom D, Drew D & de Gier, J.W (2013) Optimizing *E. coli*-based membrane protein production using Lemo21(DE3) and GFP-fusions. *Methods Mol Biol* **1033**: 381–400.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp* Ser **41**: 95-98.

He J, Sung Y, Krajmalnik-Brown R, Ritalahti KM & Löffler FE (2005) Isolation and characterization of *Dehalococcoides* sp. strain FL2, a trichloroethene (TCE)- and 1,2-dichloroethene-respiring anaerobe. *Environ Microbiol* **7**: 1442–1450.

He J Z, Ritalahti KM, Aiello MR, & Löffler FE **(**2003) Complete detoxification of vinyl chloride by an anaerobic enrichment culture and identification of the reductively dechlorinating population as a *Dehalococcoides* species. *Appl Environ Microbiol* **69:** 996–1003.

Henikoff S, Henikoff JG, Alford WJ & Pietrokovski (1995) Automated constructio12.n and graphical presentation of protein blocks from unaligned sequences. *Gene* **163**: 17–26.

Holliger C, Wohlfarth G, & Diekert G (1999) Reductive dechlorination in the energy metabolism of anaerobic bacteria. *FEMS Microbiol Rev* **22**: 383– 398.

Holliger C, Regeard C & Diekert G (2003) Dehalogenation by anaerobic bacteria. In: Häggblom, M.M., Bossert, I.D. (Eds.), Dehalogenation: *Microbial Processes and Environmental Applications. Kluwer Academic Publishing*, Boston, USA, pp. 115– 157.

Hug LA, Maphosa F, Leys D, Löffler FE, Smidt H, Edwards EA & Adrian L (2013) Overview of organohalide-respiring bacteria and a proposal for a classification system for reductive dehalogenases. *Philos T R. Soc B* **368**: 20120322.

Hazra AB, Tran JL, Crofts TS, & Taga ME (2013) Analysis of substrate specificity in CobT homologs reveal widespread preferance for DMB,the lower axial ligand of vitamin(B12).*Chem Biol* **20**: 1275-85.

Jain NK & Roy I (2009) Effect of trehalose on protein structure. *Protein Sci* **18**: 24–36.

Janssen DB (2004) Evolving haloalkane dehalogenase. *Curr Opin Chem Biol* 8: 150-9.

Jobby MK & Sharma Y (2006) Purification of a crystalline domain of *Yersinia crystalline* from inclusion bodies and its comparison to native protein from the soluble fraction. *Biomed Chromatogr* **9**: 965-63.

Johnson DR, Brodie EL, Hubbard, AE, Andersen, GL, Zinder, SH & Alvarez-Cohen L (2008) Temporal transcriptomic microarray analysis of "*Dehalococcoides ethenogenes*" strain 195 during the transition into stationary phase. *Appl Environ Microbiol* **74**: 2864-2872.

Johnson DC, Dean DR, Smith AD, Johnson MK (2005) Structure function and formation of biological iron-sulfur clusters. *Ann Rev Biochem* **74**: 247-281.

Krajmalnik-Brown R, Hölscher T, Thomson IN, Saunders FM, Ritalahti KM & Löffler FE (2004) Genetic identification of a putative vinyl chloride reductase in *Dehalococcoides* sp. strain BAV1. *Appl Environ Microbiol* **70**: 6347–6351.

Krautler B, Fieber W, Ostermann S, Fasching M, Ongania KH, Gruber K, Kratcky C, Mikl C, Siebert A and Diekert G (2003) The cofactor of tetrachloroethene reductive dehalogenase of *Dehalospirillum multivorans* is nor-pseudo B-12, a new type of natural corrinoid. *Hel Chim Acta* **86**:3698-3716.

Krogh A, Larsson B, von Heijne G & Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J Mol Biol* **305**: 567-580.

Krone UE, Laufer K, Thauer RK and Hogenkamp HPC (1989) Reductive dehalogenation of chlorinated C_1-hydrocarbons mediated by corrinoids. *Biochemistry* **28** :4908-14.

Kube M, Beck A, Zinder SH, Kuhl H, Reinhardt R & Adrian L. (2005) Genome sequence of the chlorinated compound-respiring bacterium *Dehalococcoides* species strain CBDB1. *Nat Biotechnol* **23**: 1269-1273.

Larentis AL, Argondizzo APC, Esteves GD, Jessouron E, Galler R, Medeiros MA (2011) Cloning and optimization of induction conditions for mature PsaA (pneumococcal surface adhesin A) expression in *Escherichia coli* and recombinant protein stability during long-term storage. *Protein Expr Purif* **78:** 38–47.

LaVallie ER, DiBlasio EA, Kovasis S, Grant KL, Schendel PF & McCoy JM (1993) A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E.coli* cytoplasm.*Biotechnology(N.Y.)* **2**: 187-193.

Lawrence CE, Altschul SF, Boguski MS, Liu JS, Neuwald AF & Wootton JC (1993) Detecting subtle sequence signals: a Gibbs sampling strategy for multiple alignment. *Science* **262**: 208– 214.

Lawrence JG & Roth JR 1996 (1996) Evolution of coenzyme B_{12} synthesis among enteric bacteria: evidence for loss and reacquisition of a multigene complex. *Genetics* **142**: 11-24.

Lebendiker M & Danieli T (2014) Production of prone to aggregate proteins.*FEBS letters* **588**: 236-246.

Leibly DJ, Nguyen TN, Kao LT, Hewitt SN, Barrett LK & Van Voorhis WC (2012)Stabilizing additives added during cell lysis aid in the solubilization of recombinant proteins. *PLoS One* **7**:e52482.

Lill R, Dutkiewicz R & Muelenhoff U (2006) Mechanism of iron-sulfur protein maturation in mitochondria, cytosol and nucleus of eukaryotes. *Biochimica et Biophysica Acta* **1763:** 652–667.

Löffler FE & Edwards EA (2006) Harnessing microbial activities for environmental cleanup. *Curr Opin Biotechnol* **17**: 274–284.

Löffler FE, Yan J, Ritalahti KM, Adrian L, Edwards EA, Konstantinidis KT, Müller JA, Fullerton H, Zinder SH & Spormann AM (2013) *Dehalococcoides mccartyi* gen. nov., sp. nov., obligately organohalide-respiring anaerobic bacteria relevant to halogen cycling and bioremediation, belong to a novel bacterial class, *Dehalococcoidia* classis nov., order *Dehalococcoidales* ord. nov. and family *Dehalococcoidaceae* fam. nov., within the phylum *Chloroflexi*. *Int J Syst Evol Microbiol* **63**: 625-35.

Maymó-Gatell X, Chien YT, Gossett JM, & Zinder SH **(**1997) Isolation of a bacterium that reductively dechlorinates tetrachloroethene to ethene. *Science* **276:** 1568–1571.

Maymó-Gatell, X, Nijenhuis I, &. Zinder SH **(**2001) Reductive dechlorination of *cis*-dichloroethene and vinyl chloride by "*Dehalococcoides ethenogenes*." *Environ Sci Technol* **35:** 516–521.

Mac Nelly A, Kai M, Svatos A, Diekert G & Schubert T (2014) Functional heterologous production of reductive dehalogenases from *Desulfitobacterium hafniense* strains. *Appl Environ Microbiol*.doi:10.1128/AEM.00881-14.

Magnuson JK, Romnie MF, Burris DR & Kingsley MT (2000) Trichloroethene reductive dehalogenase from *Dehalococcoides ethenogenes*: Sequence of *tceA* and substrate range characterization. *Appl Environ Microbiol* **66**: 1541-5147.

Maillard J, Schumacher W, Vazquez F, Regeard C, Hagen WR & Holliger C (2003) Characterization of the corrinoid iron-sulfur protein tetrachloroethene reductive dehalogenase of *Dehalobacter restrictus*. *Appl Environ Microbiol* **69**: 4628–4638.

Mailiard J, Genevaux P & Holliger C. (2011) Redundancy and specificity of multiple trigger factor chaperons in *Desulfitobacteria*. *Microbiology* **157**: 2410-2421.

Malkin R & Rabinowitz JC (1966) The reconstitution of clostridial ferredoxin. *Biochem Biophys Res Commun* **(6)**: 822-7.

Malakhov MP, Mattern MR, Malakhova OA, Drinker M, Weeks SD & Butt TR (2004)SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. *J Struct Funct Genomics*. **5**: 75-86.

Mancia F, Keep NH, Nakagawa A, Leadlay PF, McSweeney S, Rasmussen B, Bosecke P, Diat O & Evans PR (1996) How coenzyme B12 radicals are generated: the crystal structure of methylmalonyl-coenzyme A mutase at 2 A resolution. *Structure* **4**: 399-50.

Maness AD, Bowman KS , Yan KS, Rainey FA, & Moe WM (2012) *Dehalogenimonas spp*. can reductively dehalogenate high concentrations of 1,2-dichloroethane, 1,2-dichloropropane, and 1,1,2-trichloroethane. *AMB Express* **2**:54.

Marzorati M, De Ferra F, Van Raemdonck H, Borin S, Allifranchini E, Carpani G, Serbolisca L, Verstraete W, Boon N & Daffonchio D (2007) A novel reductive dehalogenase ,identified in a contaminated groundwater enrichment culture and in *Desulfitobacterium dichloroeliminans* strain DCA1, is linked to dehalogenation of 1,2-dichloroethane. *Appl Environ Microbiol* **73**: 2990-9.

McMurdie PJ, Behrens SF & Müller JA *et al*. (2009) Localized plasticity in the streamlined genomes of vinyl chloride respiring *Dehalococcoides. PLOS Genet* **5**: e1000714.

Miller E, Wohlfarth G & Diekert G (1998) Purification and characterization of the tetrachloroethene reductive dehalogenase of strain PCE-S. *Arch Microbiol* **169**: 497–502.

Miroux B & Walker JE (1996) Over-production of proteins in *Escherichia coli*: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol* **260**: 289–298.

Moe WM, Yan J, Nobre MF, da Costa MS & Rainey FA (2009) *Dehalogenimonas lykanthroporepellens* gen. nov., sp. nov., a reductively dehalogenating bacterium isolated from chlorinated solvent-contaminated groundwater. *Int J Syst Evol Microbiol* **59**: 2692-2697.

Mrázek, J. and Xie, S. (2006) Pattern Locator: a new tool for finding local sequence patterns in genomic DNA sequences *Bioinformatics* **22**: 3099-3100.

Müller JA, Rosner BM, Abendroth G Von, Meshulam-Simon G, McCarty PL & Spormann AM (2004) Molecular identification of the catabolic vinyl chloride reductase from *Dehalococcoides sp*. strain VS and its environmental distribution. *Appl Environ Microbiol* **70**: 4880–4888.

Nallamsetty S & Waugh DS (2006) Solubility-enhancing proteins MBP and NusA play a passive role in the folding of their fusion partners. *Protein Expr. Purif.* **45**:175–182.

Ohana RF, Encell LP, Zhao K, Simpson D, Slater MR, Urh M & Wood KV (2009) Halo Tag 7: a genetically engineered tag that enhances bacterial expression of soluble proteins and improves protein purification. *Protein Expr Purif* **68**: 110-20.

Neumann A, Wohlfarth G & Diekert G (1998) Tetrachloroethene dehalogenase from *Dehalospirillum multivorans*: cloning, sequencing of the encoding genes, and expression of the *pceA* gene in *Escherichia coli. J Bacteriol* **180**: 4140–4145.

Ni S, Fredrickson JK, & Xun L **(**1995) Purification and characterization of a novel 3-chlorobenzoate-reductive dehalogenase from the cytoplasmic membrane of *Desulfomonile tiedjei* DCB-1. *J Bacteriol* **177:** 5135–5139.

Notredame C, Higgins DG, Heringa J, 2000 T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J Mol Biol* **302**: 205– 217.

Okeke B C, Chang YC, Hatsu M, Suzuki T & K. Takamizawa **(**2001) Purification, cloning, and sequencing of an enzyme mediating the reductive dechlorination of tetrachloroethylene (PCE) from *Clostridium bifermentans* DPH-1. *Can J Microbiol* **47:**448–456.

Ou JX, Wang L, Ding XL, Du JY, Zhang Y, Chen HP,. Xu AL (2004) Stationary phase protein overproduction is a fundamental capability of *Escherichia coli*.*Biochem Biophys Res Co* **314**: 174–180.

Padilla-Crespo E, Yan J, Swift C, Wagner DD, Chourey K, Hettich RL, Ritalahti KM & Löffler FE (2014) Identification and environmental distribution of *dcpA*, which encodes the reductive dehalogenase catalyzing the dichloroelimination 1,2-dichloropropane to propene in organohalide-respiring *Chloroflexi*. *Appl Environ Microbiol* **80**: 808-818.

Park HJ, Jeon JH, Kang SG, Lee JH, Lee SA & Kim HK (2007) Functional expression and refolding of new alkaline esterase, EM2L8 from deep sea sediment metagenome. *Protein Expr Purif* **52**: 340-7.

Parthasarathy A, Stich AT, Lohner ST, Lesnefsky A, Britt R.D. & Spormann AM (2015) Biochemical and EPR-spectroscopic Investigations into Heterologously Expressed Vinyl Chloride Reductive Dehalogenase(VcrA) from *Dehalococcoides mccartyi* strain VS. *J.Am.Chem.Soc.* DOI: 10.1021/ja511653d.

Papaneophytou CP, Rinotas V, Douni E & Kontopidis G (2013) A statistical approach for optimization of RANKL overexpression in *Escherichia coli*: purification and characterization of the protein. *Protein Expr Purif* **90**: 9–19.

Papaneophytou CP & Kontopidis G (2014) Statistical approaches to maximize recombinant protein expression in *Escherichia coli*: A general review. *Protein Expr Purif* **94**: 22–32.

Payne KA, Quezada CP, Fisher K, Dunstan MS, Collins FA, Sjuts H, Levy C, Hay S, Rigby SEJ & Leys D (2015) Reductive dehalogenase structure suggests a mechanism for B12-dependent dehalogenation. *Nature* **7535**:513-516.

Pedelacq JD, Pitch E, Liong EC, Berendzen J, Kim CY, Rho BS, Park MS, Terwilliger TC &Waldo GS (2002) Engineering soluble proteins for structural genomics. *Nat Biotechnol* **20**: 927-932.

Providenti MA & Wyndham C (2001) Identification and functional characterization of CbaR, a MarR-like modulator of the *cbaABC*-encoded chlorobenzoate catabolism pathway. *Appl Environ Microbiol* **67**: 3530-3541.

Porath J (1992) Immobilized metal ion affinity chromatography. *Protein Expr Purif* **4**: 263-81.

Regeard C, Maillard J & Holliger C (2004) Development of degenerate and specific PCR primers for the detection and isolation of known and putative chloroethene reductive dehalogenase genes,. *J Microbiol Methods* **56**:107-18.

Rahm BG & Richardson RE (2008) Correlation of respiratory gene expression levels and pseudo-steady state PCE respiration rates in *Dehalococcoides ethenogenes*. *Environ Sci Technol* **42**: 416-421.

Rao PV, Bhaduri S, Jiang J & Holm HR (2004) Sulfur bridging interactions of cis-planar NiII-S2N2 coordination units with nickel(II), copper(I,II), zinc(II), and mercury(II): a library of bridging modes, including NiII(micro2-SR)2MI,II rhombs. *Inorg Chem* **19**: 5833-49.

Rowe AR, Lazar BJ, Morris RM & Richardson RE (2008) Characterization of a dechlorinating mixed culture: community structure and comparisons of gene expression in planktonic and biofloc-associated *Dehalococcoides* and *Methanospirillum*. *Appl Environ Microbiol* **74**: 6709-6719.

Rozen S & Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (Krawetz S & Misener S, eds) pp 365-386. Humana Press, Totowa, NJ.

Rose TM, Schultz ER, Henikoff JG, Pietrokovski S, McCallum CM & Henikoff S (1998) Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. *Nucleic Acids Res* **26**: 1628–1635.

Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, Lightfoot S, Menzel W, Granzow M & Ragg T (2006) The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol* **7**:3.

Schein CH& Noteborn MHM (1988) Formation of soluble recombinant proteins in *Escherichia coli* is favored by lower growth temperature. *Nat Biotechnol* **6**: 291–294.

Schein CH (1989) Production of soluble recombinant proteins in bacteria. *Bio/technology (New York, N. Y. 1983)* **7**: 1141–1149.

Schlapschy M & Skerra A (2011) Periplasmic chaperones used to enhance functional secretion of proteins in *E.coli*. *Methods Mol Biol* **705**: 2011-24.

Seshadri R, Adrian L & Fouts DE *et al*. (2005) Genome sequence of the PCE-dechlorinating bacterium *Dehalococcoides ethenogenes*. *Science* **307**: 105-108.

Shigeyuki Yokoyama (2003) Protein expression systems for structural genomics and proteomics. *Current opinion in chemical biology* **7**: 39-43.

Shaw MK& Ingraham JL (1967) Synthesis of macromolecules by *Escherichia coli* near the minimal temperature for growth. *J Bacteriol* **94**: 157–164.

Shethna YI, Wison PW, Hansen RE & Beinert H (1964) Identification by isotropic substitution of the EPR signal at g= 1.94 in a non-heme iron protein from *Azotobacter*. *Proc Natl Acad Sci U S A.* **52(5)**: 1263–1271.

Shey J & van der Donk WA (2000) Mechanistic studies on the Vitamin B_{12}-Catalyzed Dechlorination of Chlorinated Alkenes. *J Am Chem Soc* **49**: 12403-12404.

Siddaramappa S, Challacombe JF & Delano SF *et al.* (2012) Complete genome sequence of *Dehalogenimonas lykanthroporepellens* type strain (BL-DC-9T) and comparison to "*Dehalococcoides*" strains. *Standards in Genomic Sci* **6**: 251-264.

Singh K, Mehta G & Chhatpar HS (2009) Optimization of medium constituents for improved chitinase production by *Paenibacillus sp* D1 using statistical approach. *Lett Appl Microbiol* **49**: 708–714.

Singh SS & Panda AK (2005) Solubilization and refolding of bacterial inclusion body proteins. *J Biosc and Bioengg* **99**: 303-310.

Song JA, Lee DS, Park JS, Han KY & Lee J (2012) The N-domain of *Escherichia coli* phosphoglycerate kinase is a novel fusion partner to express aggregation-prone heterologous proteins. *Biotechnol Bioeng* **109**: 325–335.

Song JM, An YJ, Kang MH, Lee YH& Cha SS (2012) Cultivation at 6–10 _C is an effective strategy to overcome the insolubility of recombinant proteins in *Escherichia coli*. *Protein Expr Purif* **82**: 297-301.

Song H, Li G, Mai W, Huang G, Chen K, Zhou Y & Chen H (2014) Codon optimization enhances protein expression of *Bombyx mori* nucleopolyhedrovirus DNA polymerase in *E. coli*. *Curr Microbiol* **68**: 293-300.

Smidt H, & de Vos WM (2004) Anaerobic microbial dehalogenation. *Annu Rev Microbiol* **58**: 43-73.

Smith HO, Annau TM & Chandrasegaran S (1990) Finding sequence motifs in groups of functionally related proteins. *Proc Natl Acad Sci* **87**: 826– 830.

Smith DB & Johnson KS (1988) Single-step purification of polypeptides expressed in *Escherechia coli* as fusions with glutathione S-transferase. *Gene* **67**: 31–40.

Sreenath HK, Bingman CA, Buchan BW, Seder KD, Burns BT, Geetha HV, Jeon WB, Vojtik FC, Aceti DJ & Frederick RO(2005) Protocols for production of selenomethionine-labeled proteins in 2-Lpolyethylene terephthalate bottles using auto-induction medium, *Protein Expr Purif* **40**: 256–267.

Stanley NR, Palmer T & Berks BC (2000) The twin arginine consensus motif of Tat signal peptides is involved in Sec-independent protein targeting in *Escherichia coli*. *J Biol Chem* **275**: 11591-11596.

Suyama A, Yamashita M, Yoshino S & Furukawa K (2002) Molecular characterization of the pceA reductive dehalogenase of *Desulfitobacterium* sp. strain Y51. *J Bacteriol* **184**: 3419–3425.

Sjuts H, Fisher K, Dunstan MS, Rigby SE & Leys D (2012) Heterologous expression, purification and cofactor reconstitution of the reductive dehalogenase PceA from *Dehalobacter restrictus*. *Protein Expr Purif* **85**: 224–229.

Tang S, Chan WW M, Fletcher KE, Seifert J, Liang X, Löffler FE, Edwards EA & Adrian L (2013) Functional characterization of reductive dehalogenases by using blue native polyacrylamide gel electrophoresis. *Appl Environ Microbiol* **79**: 974-981.

Tang S & Edwards EA (2013) Identification of *Dehalobacter* reductive dehalogenases that catalyse dechlorination of chloroform, 1, 1,1-trichloroethane and 1,1-dichloroethane. *Philos T R Soc B* **368**: 20120318 .

Tabor S & Richardson CC (1985) A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc Natl Acad Sci USA* **82**: 1074–1078.

Tabandeh F, Khodabandeh M, Yakhchali B, Habib-Ghomi H & P. Shariati (2008) Response surface methodology for optimizing the induction conditions of recombinant interferon beta during high cell density culture. *Chem Eng Sci* **63**: 2477–2483.

Tsumoto K, Abe R, Ejima D & Arakawa T (2010) Non denaturing solubilization of inclusion bodies. *Curr Pharm Biotechnol* **11**: 309-12.

Vasina JA & Baneyx F (1996) Recombinant protein expression at low temperatures under the transcriptional control of the major *Escherichia coli* cold shock promoter cspA. *Appl Environ Microbiol* **62**: 1444–1447.

Wagner A, Adrian L, Kleinsteuber S, Andreesen JR & Lechner, U (2009) Transcription analysis of genes encoding homologues of reductive dehalogenases in "*Dehalococcoides*" sp. strain CBDB1 by using terminal restriction fragment length polymorphism and quantitative PCR. *Appl Environ Microbiol* **75**: 1876-1884.

Wagner A, L Segler, S Kleinsteuber, G Sawers, H Smidt & U Lechner (2013) Regulation of reductive dehalogenase gene transcription in *Dehalococcoides mccartyi*. *Philos T R. Soc* B **368**: 20120317.

Waller AS, Krajmalnik-Brown R, Löffler FE & Edwards EA (2005) Multiple reductive-dehalogenase-homologous genes are simultaneously transcribed during dechlorination by *Dehalococcoides*-containing cultures. *Appl Environ Microbiol* **71**: 8257-8264.

Waldo GS, Standish BM, Berendzen J &Terwilliger TC (1999) Rapid protein folding assay using green fluoroscent protein. *Nat Biotechnol* **17**: 691-695.

Wilkinson SP & Grove A (2006) Ligand-responsive transcriptional regulation by members of the MarR family of winged helix proteins. *Curr Issues Mol Biol* **8**: 51-62.

Wyborski DL, Bauer JC, Zheng CF, Felts K, Vaillancourt P (1999) An *Escherichia coli* expression vector that allows recovery of proteins with native N-termini from purified calmodulin-binding peptide fusions. *Protein Expr Purif* **16:** 1–10.

Yan J, Rash BA, Rainey FA & WM Moe (2009) Isolation of novel bacteria within the *Chloroflex*i capable of reductive dechlorination of 1,2,3-trichloropropane. *Environ Microbiol* **11**: 833-843.

Villemur R, Saucier M, Gauthier A, & Beaudet R **(**2002) Occurrence of several genes encoding putative reductive dehalogenases in *Desulfitobacterium hafniense/frappieri* and *Dehalococcoides ethenogenes*. *Can J Microbiol* **48:** 697–706.

Yokoyama S (2003) Protein expression system for structural genomics and proteomics. *Current opinion in chemical biology* **7**: 39-43.

Wilkinson SP & Grove A (2006) Ligand-responsive transcriptional regulation by members of the MarR family of winged helix proteins. *Curr Issues Mol Biol* **8**: 51-62.

Zhou P, Lugovskoy AA & Wagner G (2001) A solubility enhancement tag (SET) for NMR studies of poorly behaving proteins. *J Bimol NMR* **20**: 11-14.

Appendix A: Results from Bioanalyzer tests to assess RNA quality

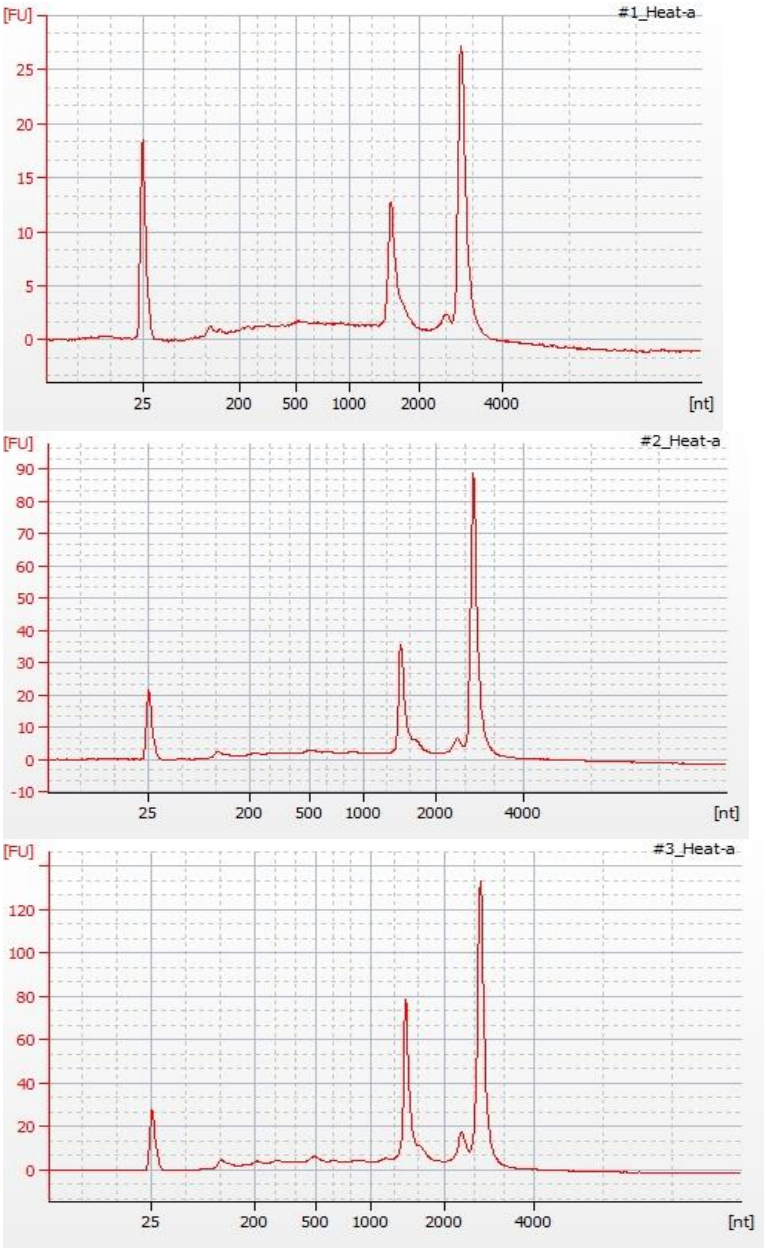


Fig A1. Electropherogram results for RNA extracted from *D. lykanthroporepellens* BL-DC-9^{T} following growth on 1,2,3-TCP (Replicates A, B, and C shown from top to bottom).

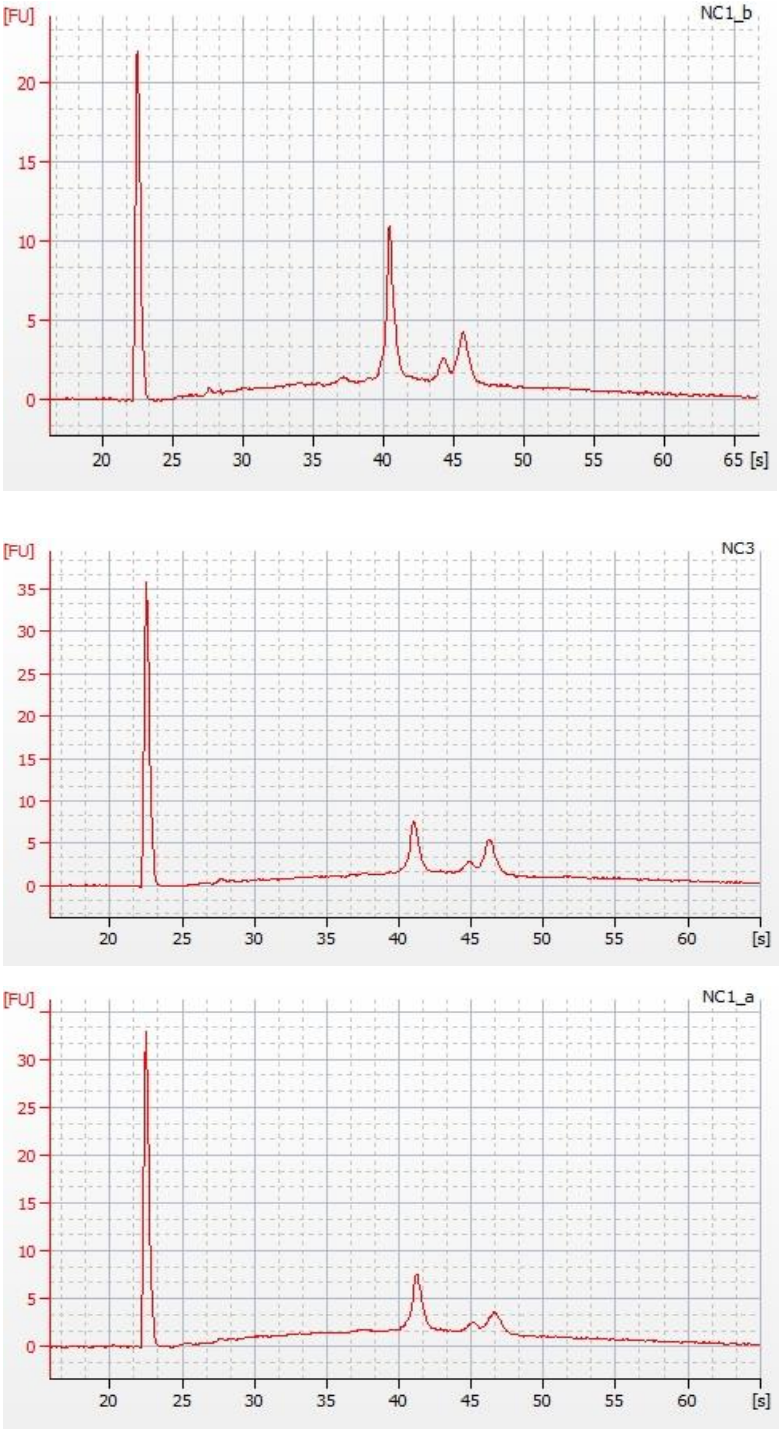


Fig A2. Electropherogram results for RNA extracted from *D. lykanthroporepellens* BL-DC-9^{T} following incubation with no solvent (controls in the 1,2,3-TCP experiment). Replicates A, B, and C shown from top to bottom.

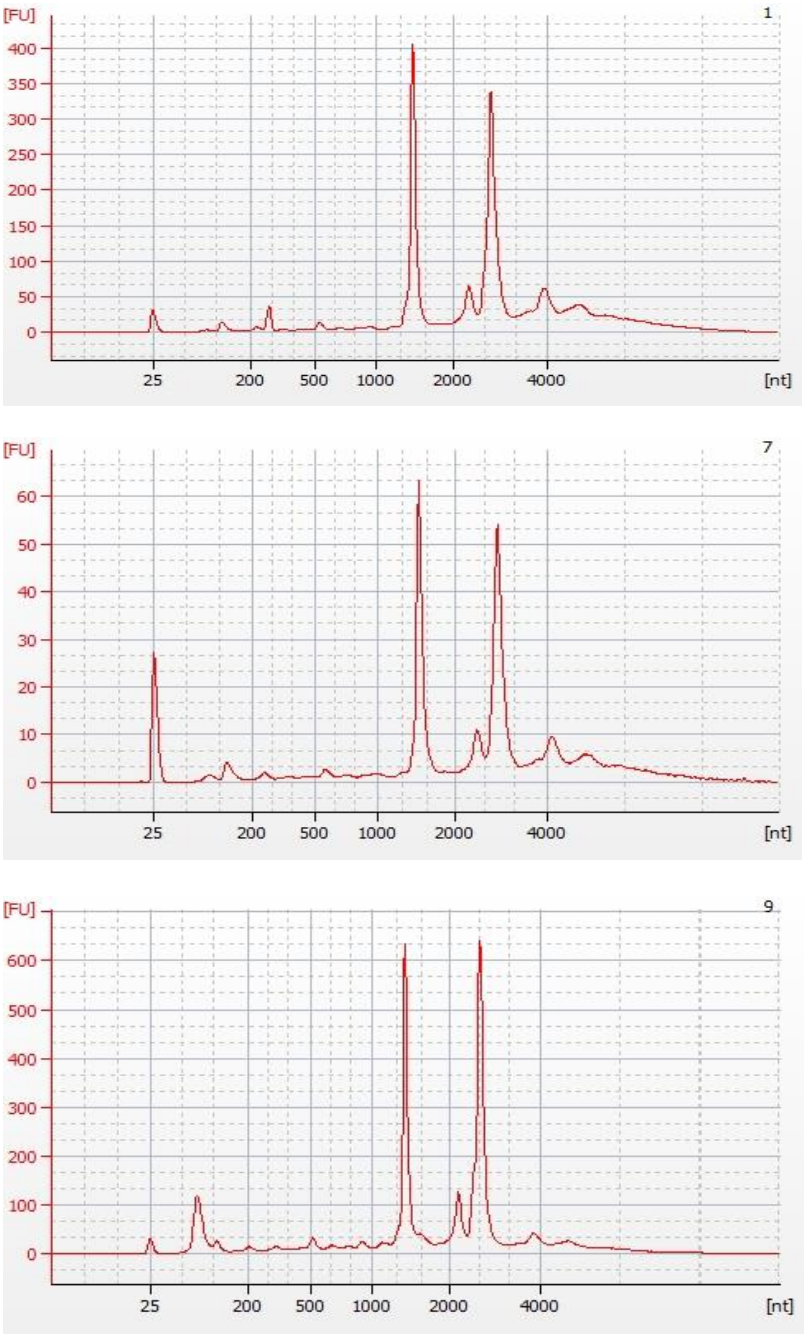


Fig A3. Electropherogram results for RNA extracted from *D. lykanthroporepellens* BL-DC-9^{T} following growth on 1,2-DCP (Replicates A, B, and C shown from top to bottom).

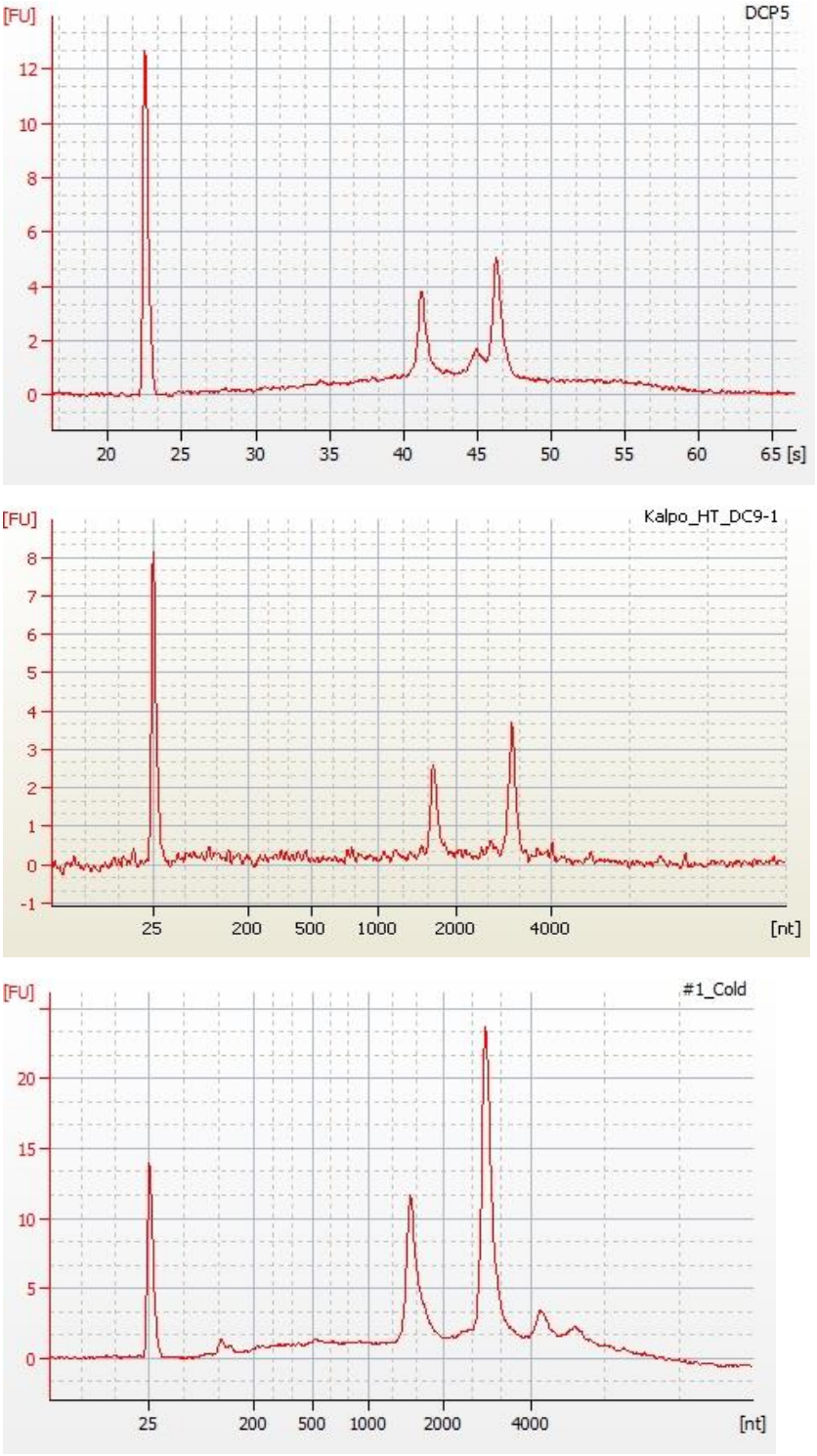


Fig A4. Electropherogram results for RNA extracted from *D. lykanthroporepellens* BL-DC-9^{T} following incubation with no solvent (controls in the 1,2-DCP experiment) (Replicates A, B, and C shown from top to bottom).

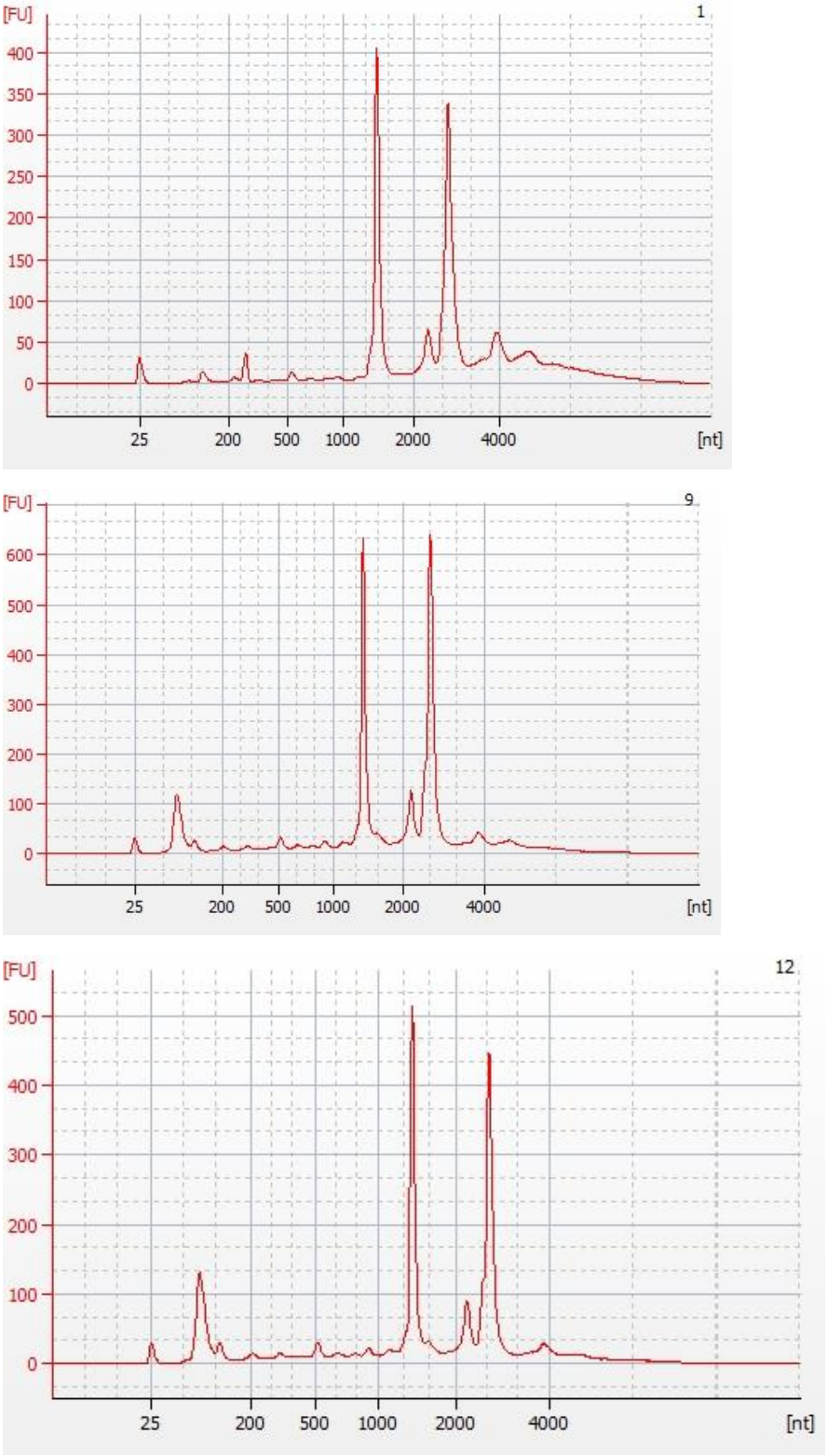


Fig A5. Electropherogram results for RNA extracted from *D. lykanthroporepellens* BL-DC-9^{T} following growth on 1,2-DCA (Replicates A, B, and C shown from top to bottom).

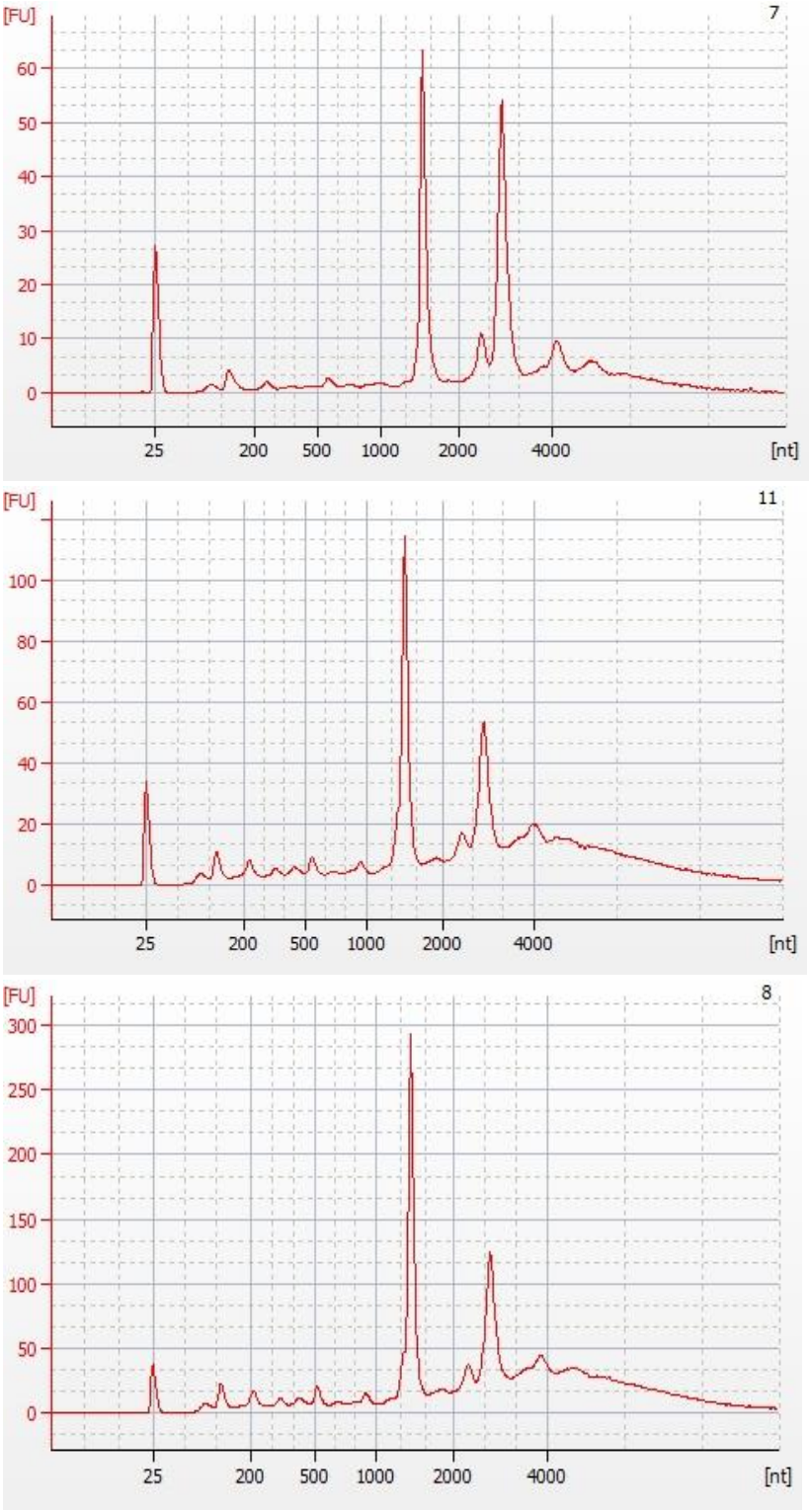


Fig A6. Electropherogram results for RNA extracted from *D. lykanthroporepellens* BL-DC-9^{T} following incubation with no solvent (controls in the 1,2-DCA experiment). Replicates A, B, and C shown from top to bottom.

Appendix B: Gel images supporting results reported in Tables 2.6 and 2.4 (PCR amplicons from all solvent replicates run on the same gel for a single primer set)

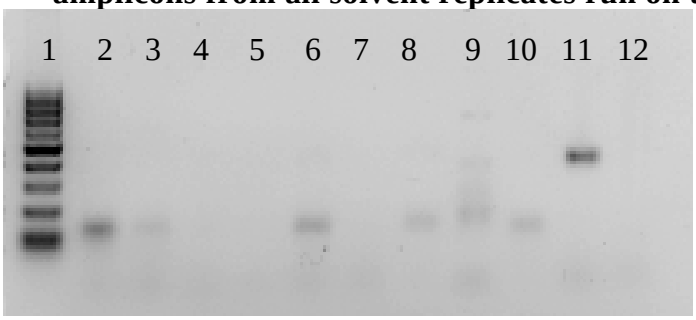


Fig B1. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_0068 (expected amplicon size 434 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

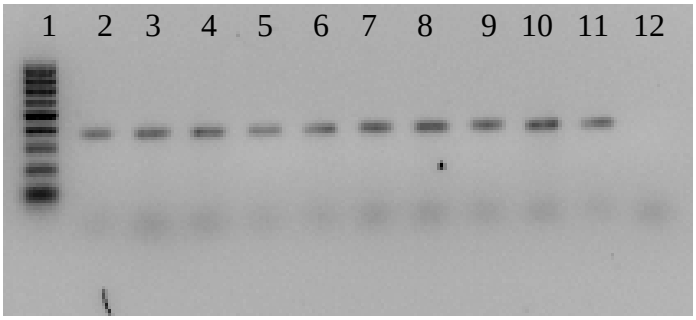


Fig B2. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_0069 (expected amplicon size 360 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

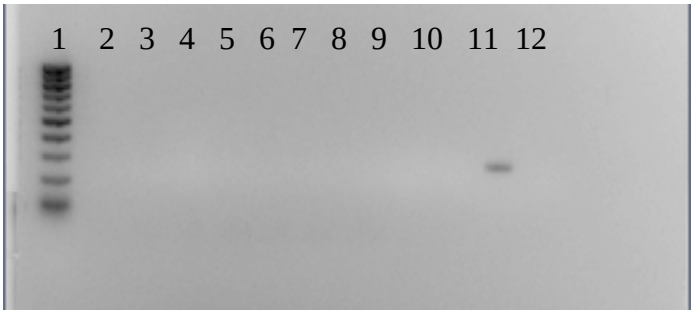


Fig B3. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_0075 (expected amplicon size 270 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

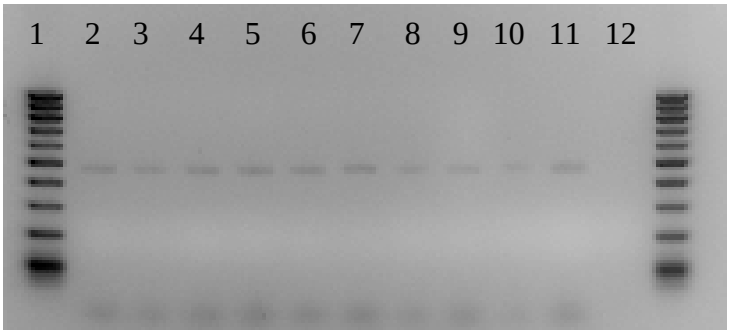


Fig B4. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_0121 (expected amplicon size 462 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

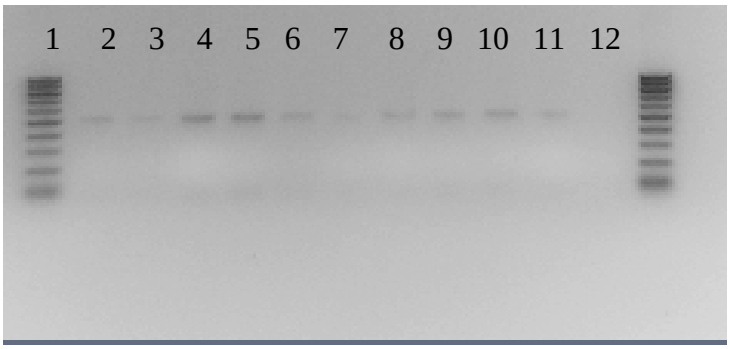


Fig B5. Gel images showing amplicons from RT-PCR employing primers targeting *rdhAB* Dhly_0156-Dhly_0157 (expected amplicon size 503 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

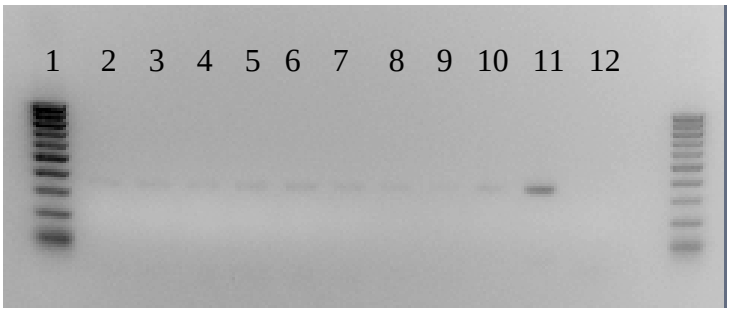


Fig B6. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_0274 (expected amplicon size 345 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

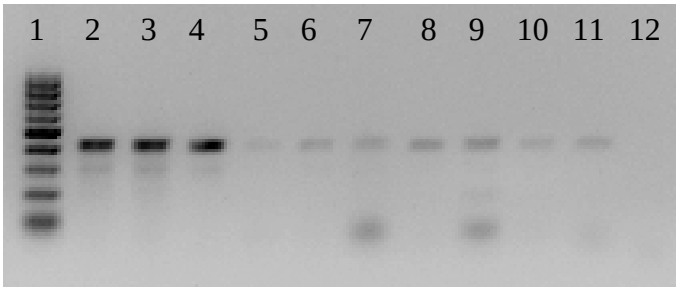


Fig B7. Gel images showing amplicons from RT-PCR employing primers targeting *rdhAB* Dhly_0275-Dhly_0276 (expected amplicon size 419 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

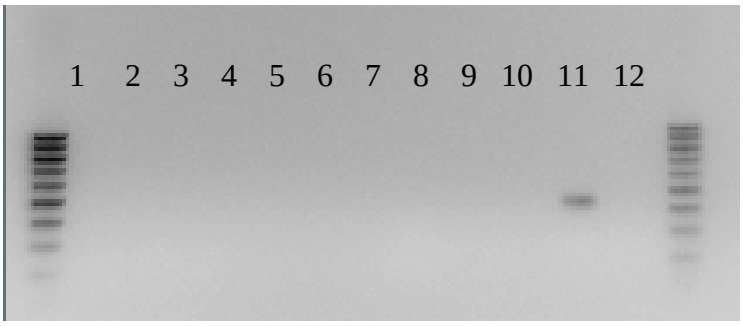


Fig B8. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_0283 (expected amplicon size 438 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

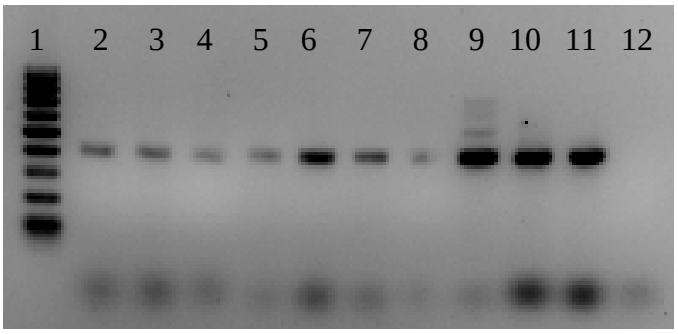


Fig B9. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_0479 (expected amplicon size 410 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

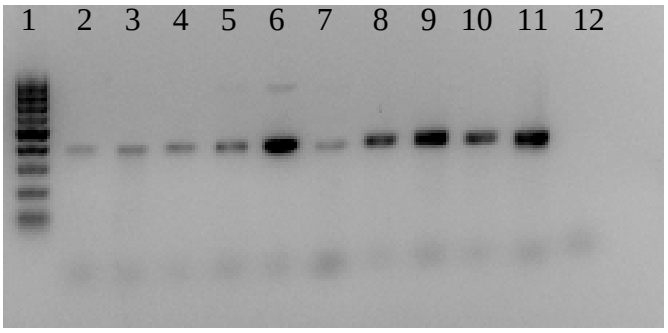


Fig B10. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_0849 (expected amplicon size 372 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

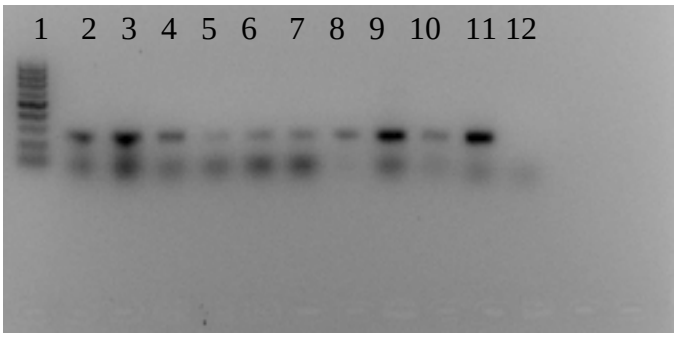


Fig B11. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_0910 (expected amplicon size 244 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

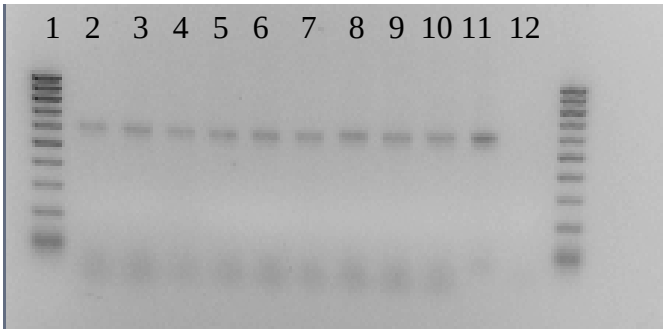


Fig B12. Gel images showing amplicons from RT-PCR employing primers targeting *rdhAB* Dhly_1053-Dhly_1054 (expected amplicon size 608 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

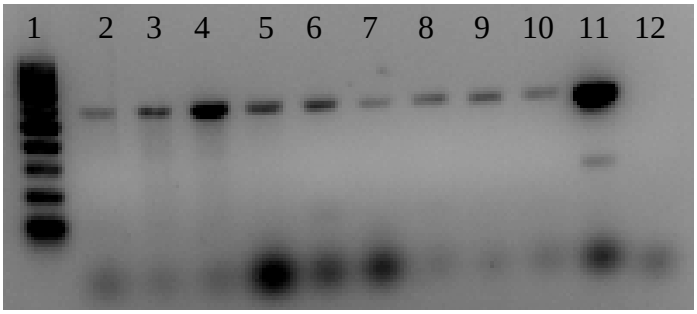


Fig B13. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_1148 (expected amplicon size 551 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

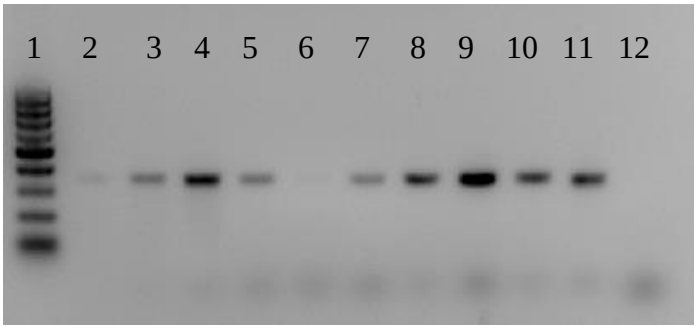


Fig B14. Gel images showing amplicons from RT-PCR employing primers targeting *rdhAB* Dhly_1151-Dhly_1152 (expected amplicon size 349 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

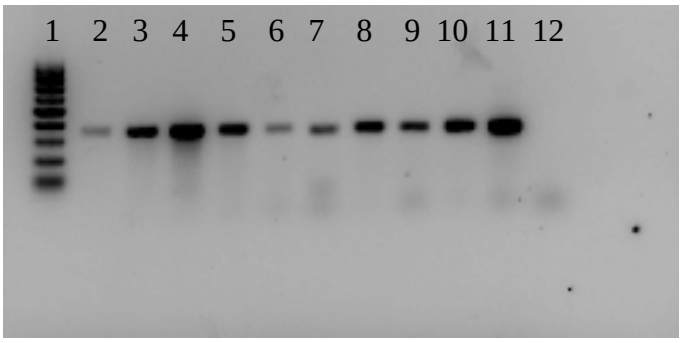


Fig B15. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_1328 (expected amplicon size 347 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

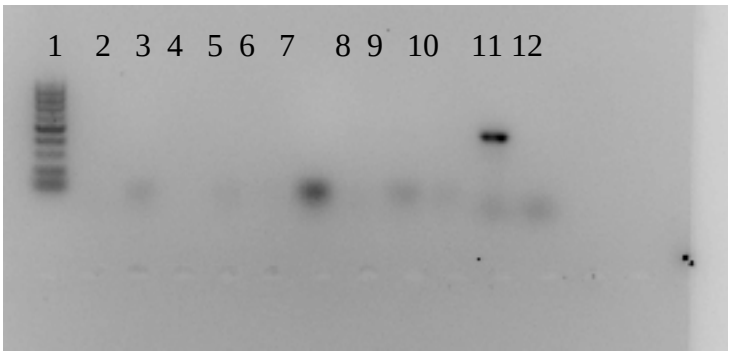


Fig B16. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_1355 (expected amplicon size 464 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

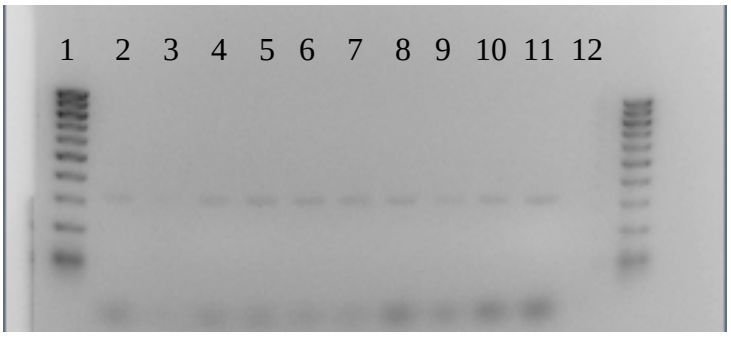


Fig B17. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_1514 (expected amplicon size 310 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

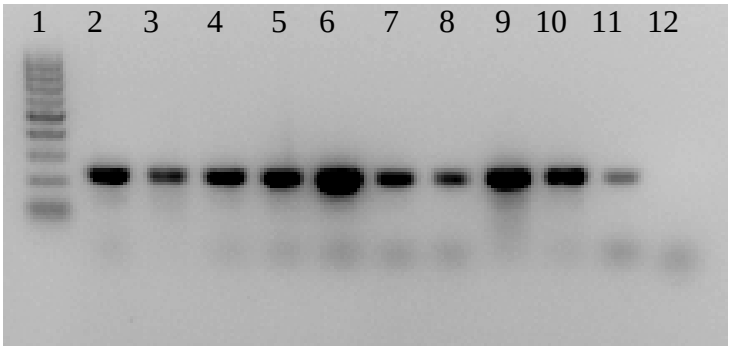


Fig B18. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_1520 (expected amplicon size 217 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

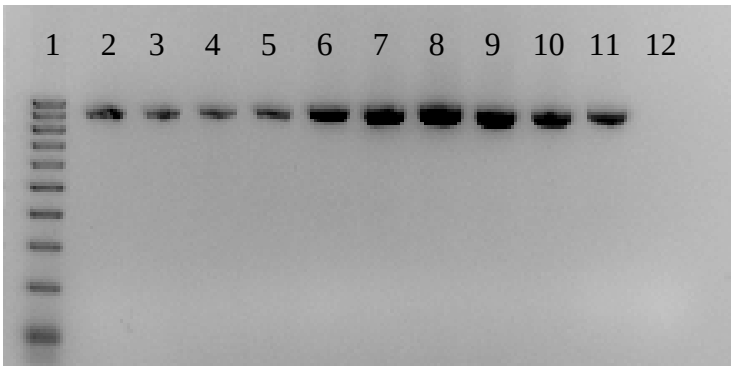


Fig B19. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_1523 (expected amplicon size 892 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

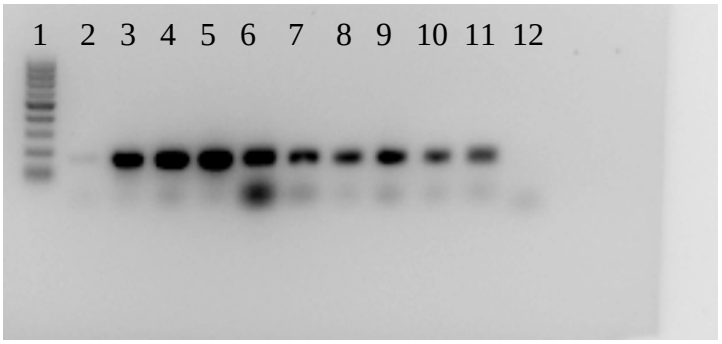


Fig B20. Gel images showing amplicons from RT-PCR employing primers targeting *rdhAB* Dhly_1524-Dhly_1525 (expected amplicon size 168 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

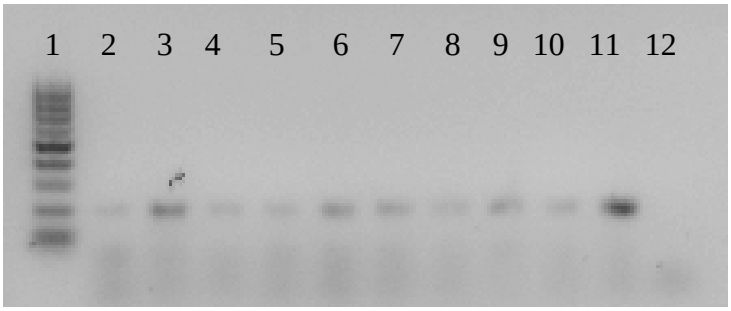


Fig B21. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_1530 (expected amplicon size 196 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

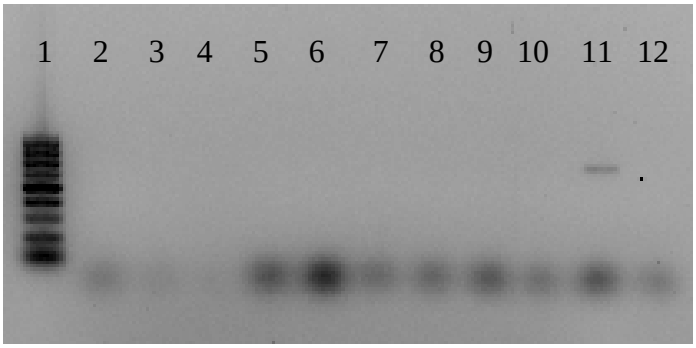


Fig B22. Gel images showing amplicons from RT-PCR employing primers targeting *rdhAB* Dhly_1533-Dhly_1534 (expected amplicon size 677 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

Fig B23. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_1540 (expected amplicon size 387 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

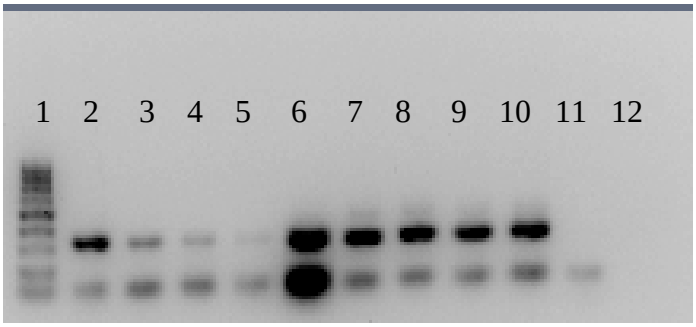


Fig B24. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_1541 (expected amplicon size 115 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

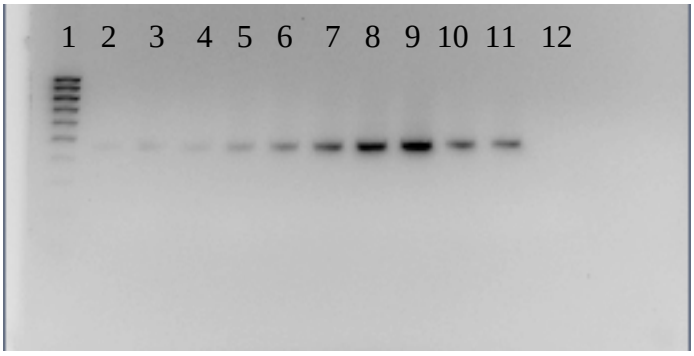


Fig B25. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_1582 (expected amplicon size 437 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

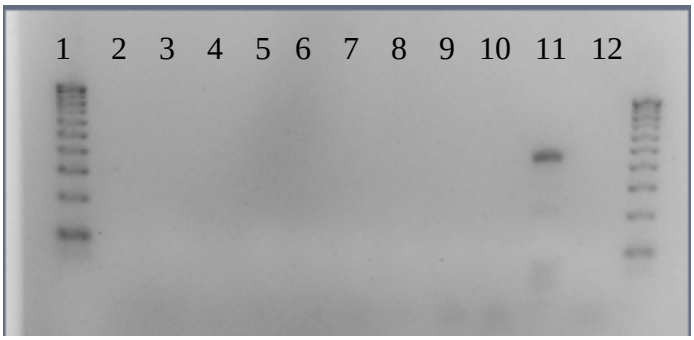


Fig B26. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_0068 (expected amplicon size 434 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

Fig B27. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_0069 (expected amplicon size 360 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

Fig B28. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_0075 (expected amplicon size 270 bp) for *D. lykanthroporepellens* BL-DC-9^{T} following growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

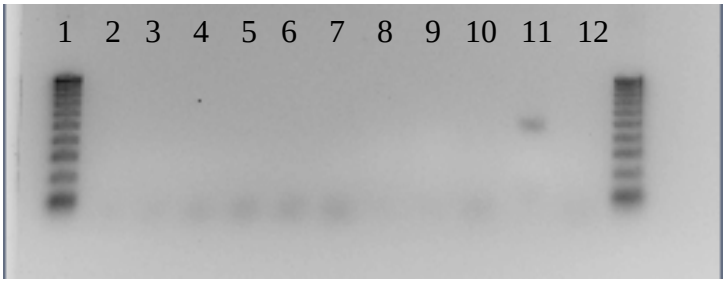


Fig B29. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_0121 (expected amplicon size 462 bp) for *D. lykanthroporepellens* BL-DC-9^{T} following growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

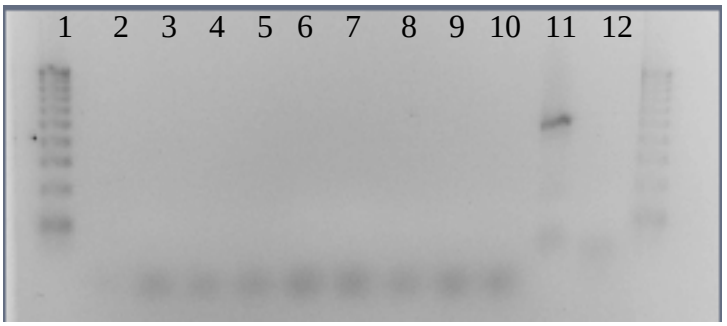


Fig B30. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_0156-Dehly_0157 (expected amplicon size 503 bp) for *D. lykanthroporepellens* BL-DC-9^{T} following growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

Fig B31. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_0274 (expected amplicon size 345 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

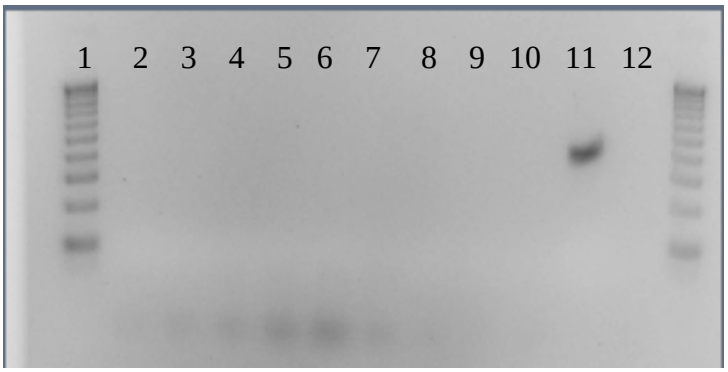


Fig B32. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_0275-Dehly 0276 (expected amplicon size 419 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

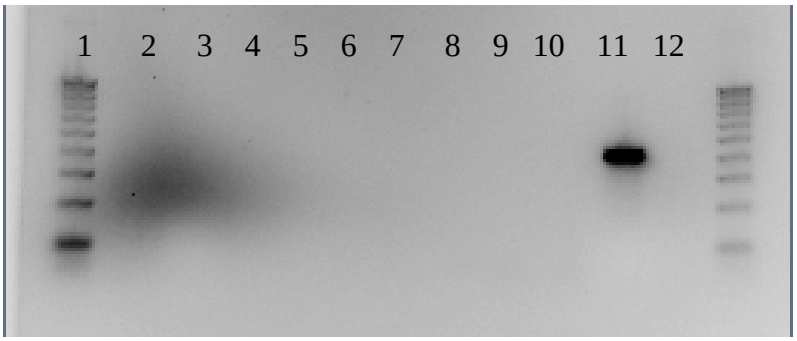


Fig B33. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_0283 (expected amplicon size 438 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

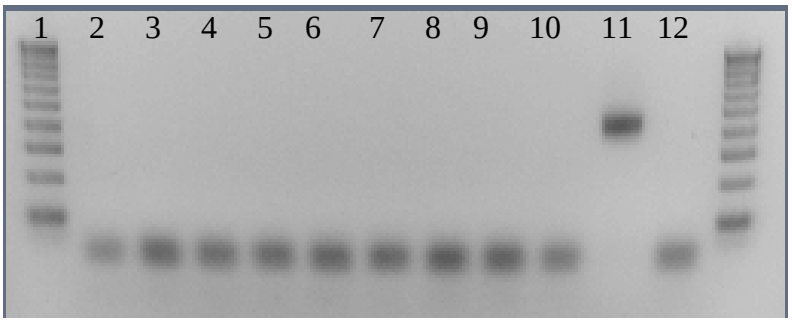


Fig B34. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_0479 (expected amplicon size 410 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

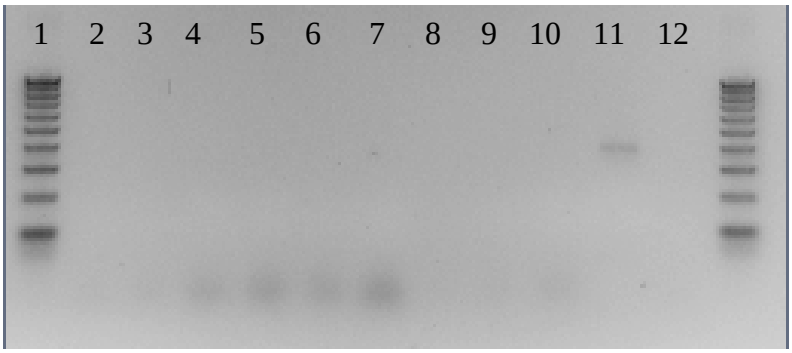


Fig B35. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_0849 (expected amplicon size 372 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

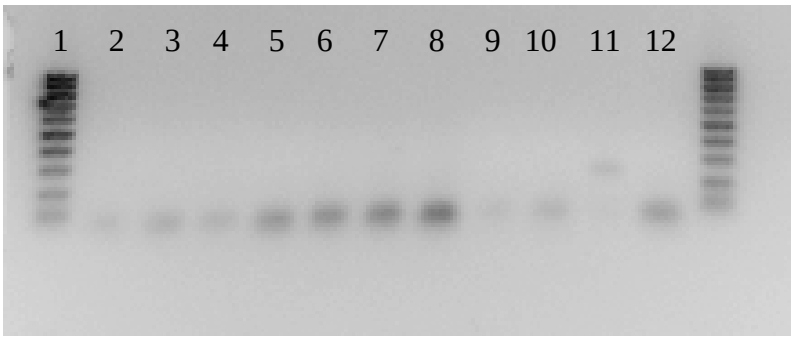


Fig B36. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_0910 (expected amplicon size 244 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

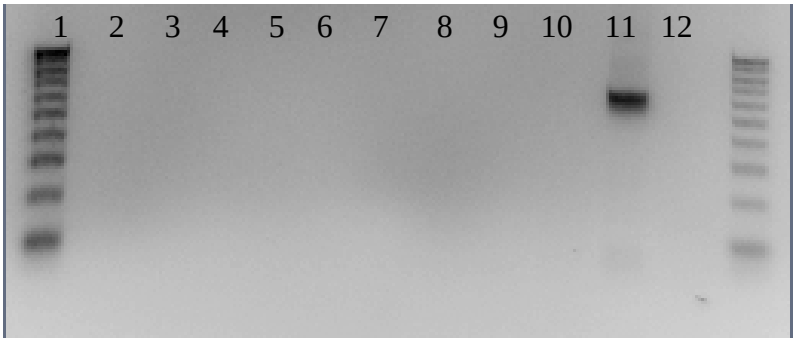


Fig B37. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_1054_Dehly_1053 (expected amplicon size 608 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

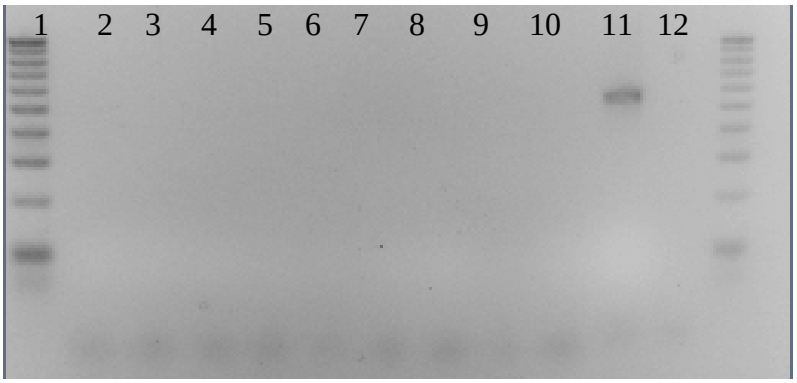


Fig B38. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_1148 (expected amplicon size 551 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

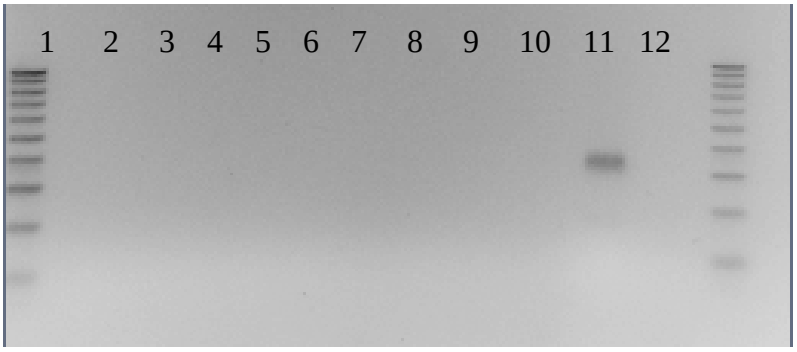


Fig B39. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_1151-Dehly-1152 (expected amplicon size 349 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

Fig B40. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_1328 (expected amplicon size 347 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

Fig B41. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_1355 (expected amplicon size 464 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

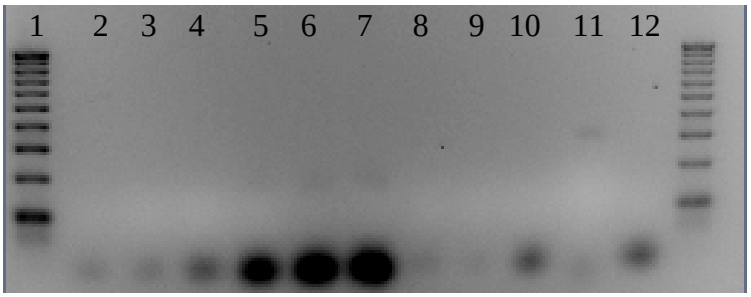


Fig B42. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_1514 (expected amplicon size 310 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

Fig B43. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_1520 (expected amplicon size 217 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

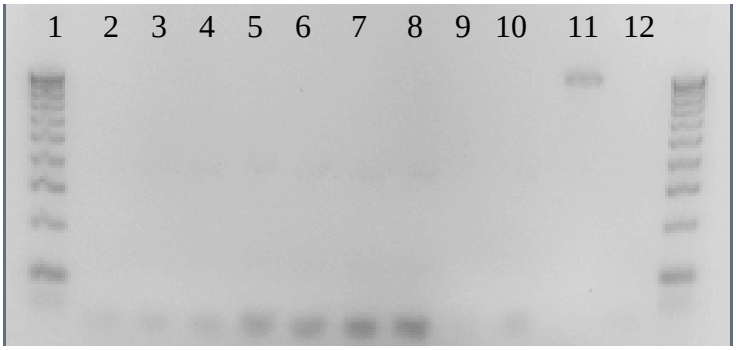


Fig B44. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_1523 (expected amplicon size 892 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth in

absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

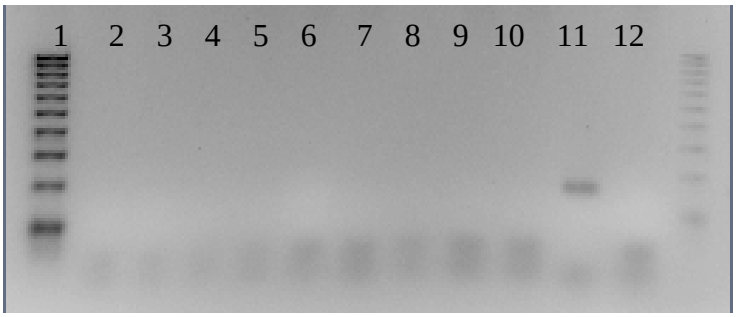


Fig B45. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_1524-Dehly_1525 (expected amplicon size 168 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

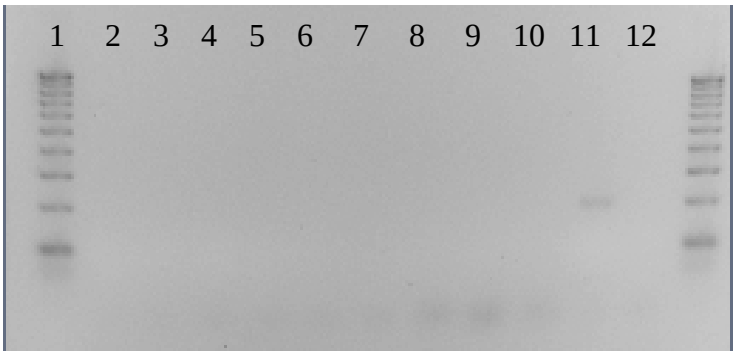


Fig B46. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_1530 (expected amplicon size 196 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

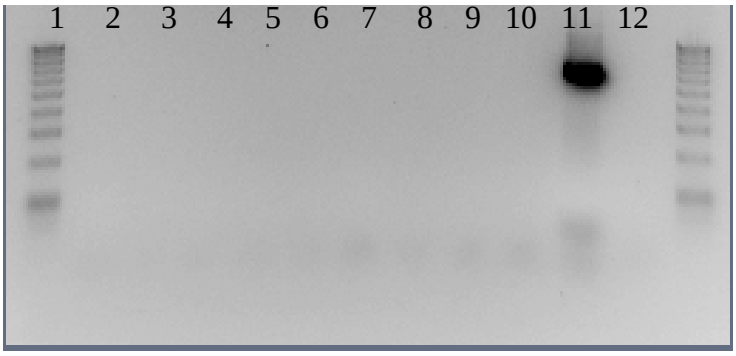


Fig B47. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_1533-Dehly_1534 (expected amplicon size 677 bp) for *D. lykanthroporepellens* BL-DC-9^T following

growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

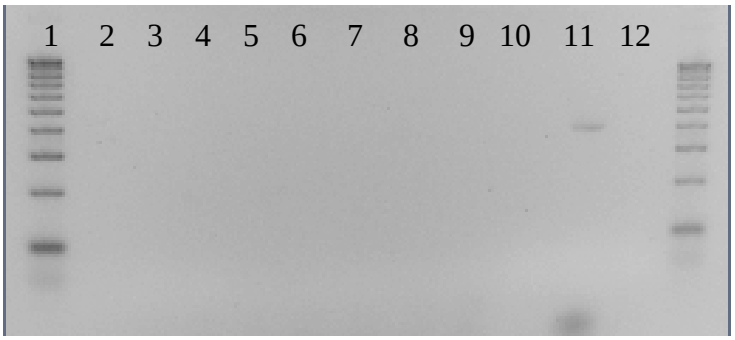


Fig B48. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_1540 (expected amplicon size 387 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

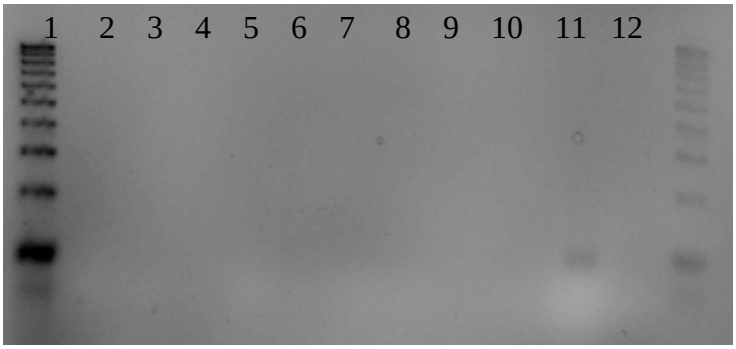


Fig B49. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_1541 (expected amplicon size 115 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

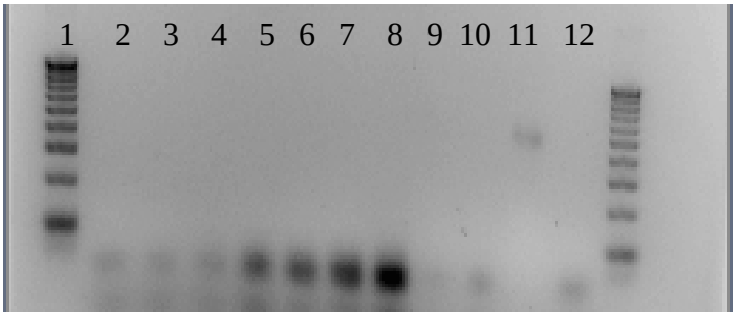


Fig B50. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_1582 (expected amplicon size 437 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

Appendix C: Gel images supporting results reported in Tables 2.5 (PCR products all solvent replicates run on the same gel for a single primer set)

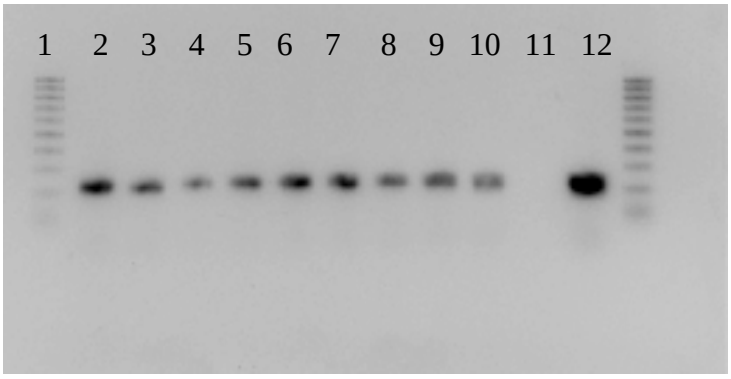


Fig C1. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_0120 (expected amplicon size 225 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 12 positive control. Lane 11 negative control.

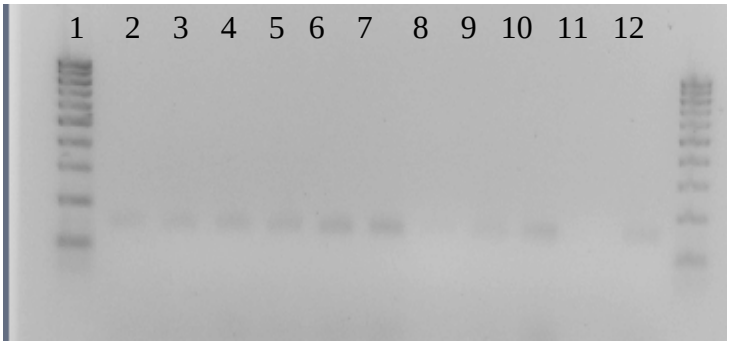


Fig C2. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_0270 (expected amplicon size 143 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 12 positive control. Lane 11 negative control.

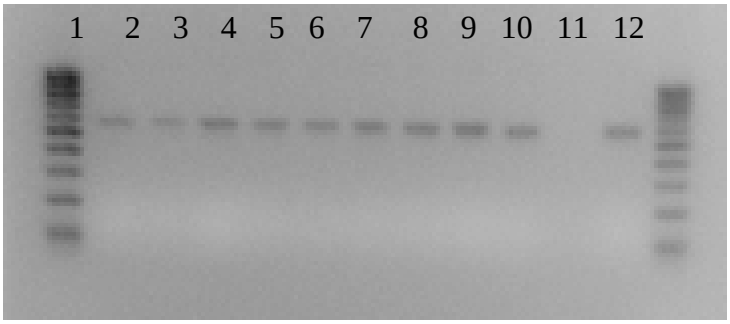


Fig C3. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_1149 (expected amplicon size 622 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 12 positive control. Lane 11 negative control.

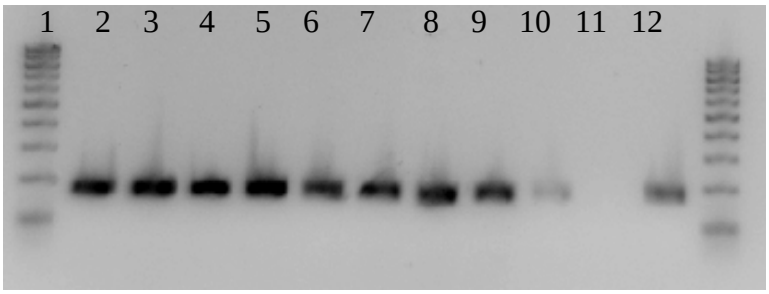


Fig C4. Gel images showing amplicons from RT-PCR employing primers targeting Dhly_1531 (expected amplicon size 185 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth on 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 12 positive control. Lane 11 negative control.

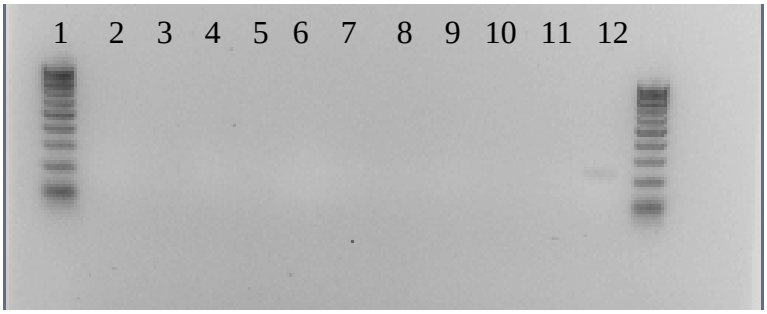


Fig C5. Gel images showing amplicons from RT-PCR employing primers targeting Dehly_0120 (expected amplicon size 225 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 12 positive control. Lane 11 negative control.

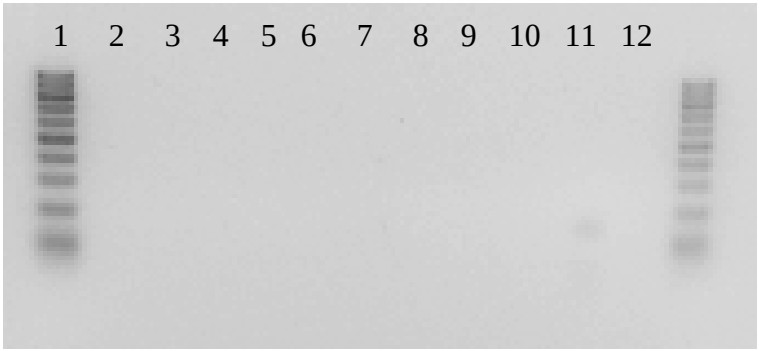


Fig C6. Gel images showing amplicons from RT-PCR employing primers targeting Dehly_0270 (expected amplicon size 143 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 11 positive control. Lane 12 negative control.

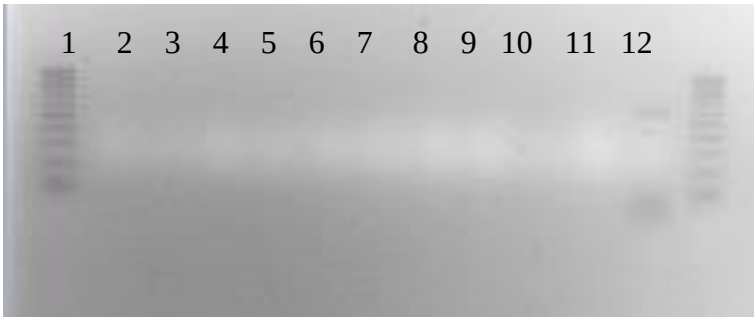


Fig C7. Gel images showing amplicons from RT-PCR employing primers targeting Dehly_1149 (expected amplicon size 622 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 12 positive control. Lane 11 negative control.

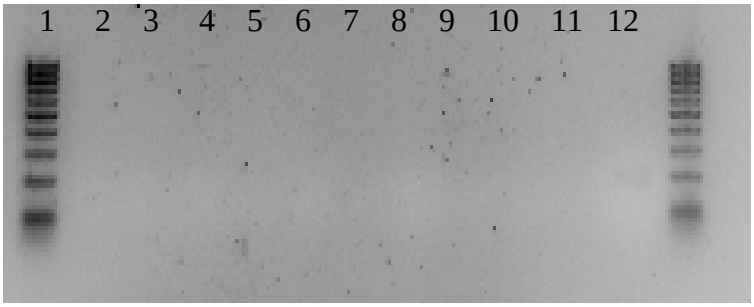


Fig C8. Gel images showing amplicons from RT-PCR employing primers targeting Dehly_1531 (expected amplicon size 185 bp) for *D. lykanthroporepellens* BL-DC-9^T following growth in absence of 1,2-DCA (lanes 2-4), 1,2-DCP (lanes 5-7), and 1,2,3-TCP (lanes 8-10). Lane 1: marker (100 bp Sigma). Lane 12 positive control. Lane 11 negative control.

Appendix D1: Gel images supporting results reported in chapter 3

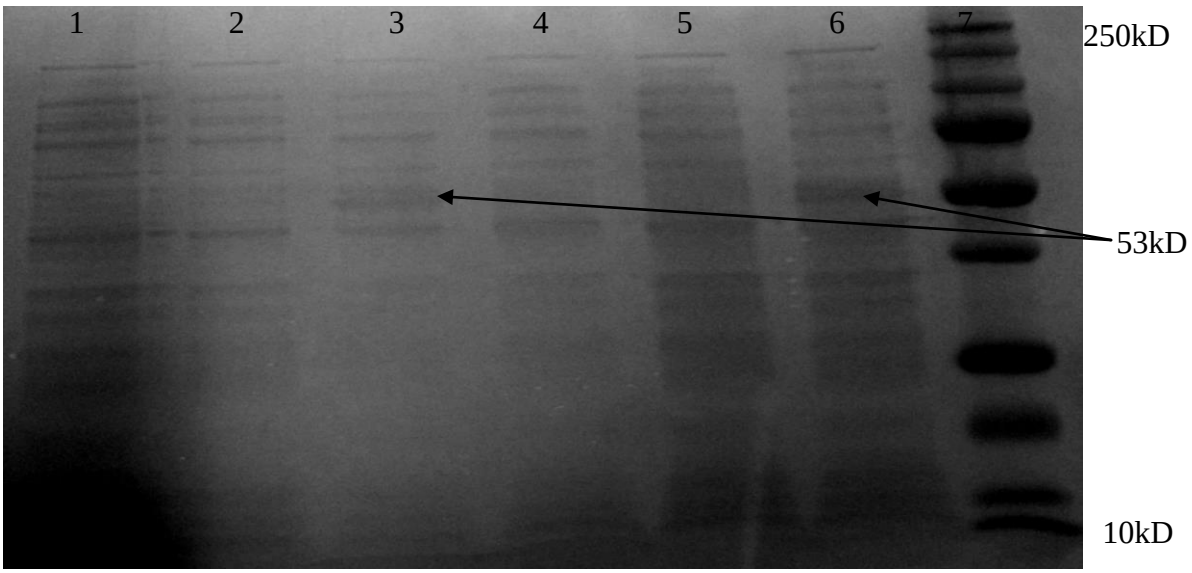


Fig D1. SDS PAGE analysis of transformed *E. coli* induced at different temperatures. Lanes 2, 3 and 6: the lysates from cells induced at temperatures of 18 °C (lane 2), 23 °C (lane 3), and 37°C (lane 6). Lanes 1, 4 & 5: the soluble fractions recovered from cells induced at temperatures of 18 °C (lane 1), 23 °C (lane 4), and 37°C (lane 5). Lane 7: Precision Plus protein standard (BioRad).

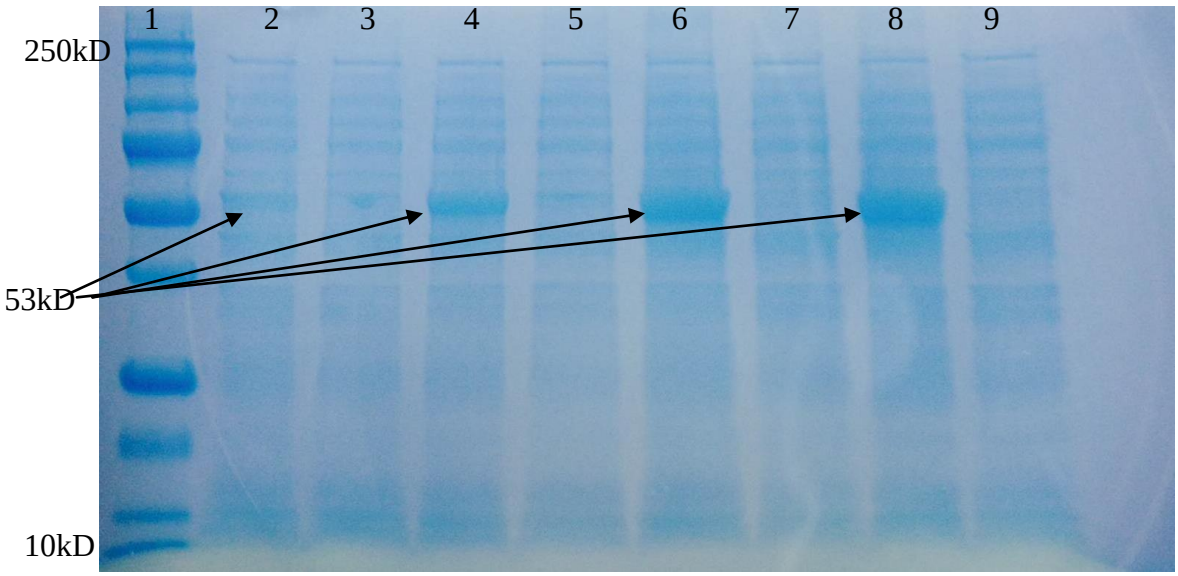


Fig D2. SDS PAGE analysis of the transformed *E. coli* cells induced for different time intervals. Lanes 2, 4, 6 and 8: cell lysate generated after induction for 0.5 hours (lane 2), 1 Hour (lane 4), 1.5 Hour (lane 6) and 2 hours (lane 8). Lanes 3, 5, 7 and 9: soluble fractions of cell lysate recovered from cells induced for 0.5 Hour (lanes 3), 1 hour (lane 5), 1.5 Hours (lane 7) and 2 hours (lane 9). Lane 1: Precision Plus protein standard (BioRad).

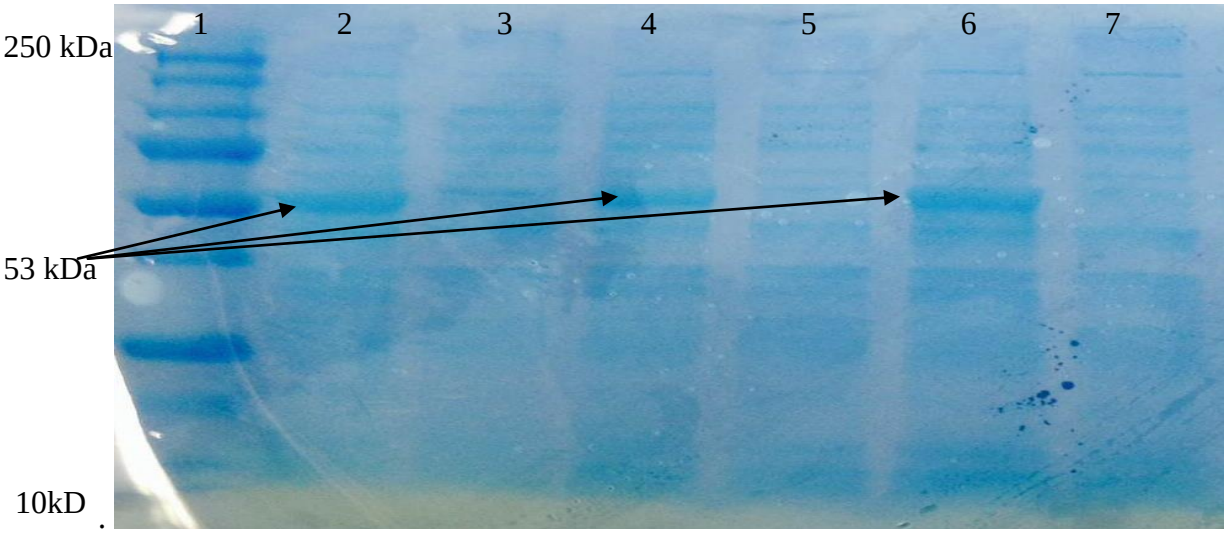
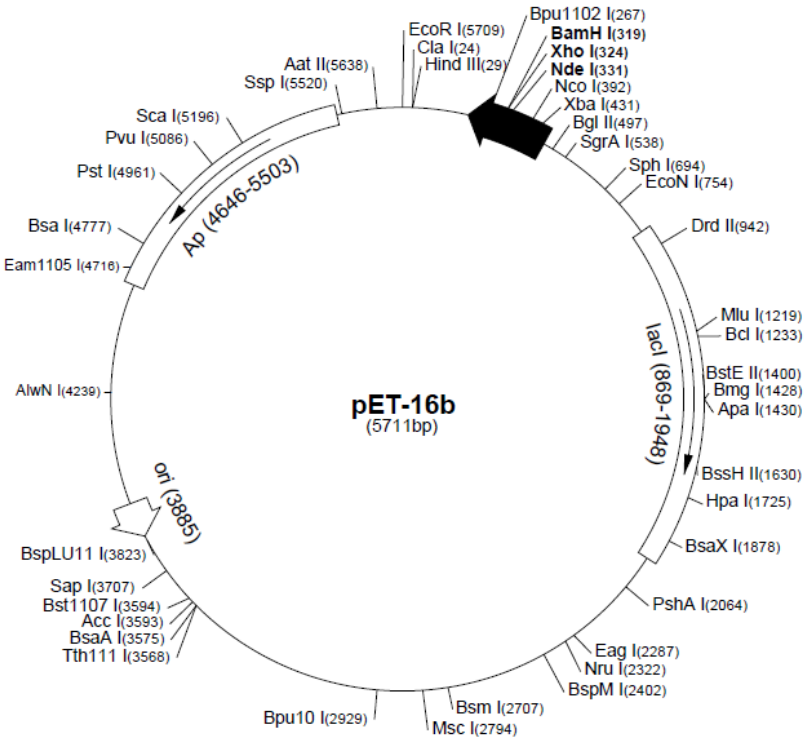


Fig D3. SDS PAGE analysis of the transformed *E. coli* cells that were induced at different IPTG concentrations. Lanes 2, 4, and 6: cell lysate generated at IPTG concentration 0.2 mM (lane 2), 0.4 mM (lane 4) and 0.6 mM (lane 6). Lanes 3, 5 and 7: soluble fractions of the cell lysate recovered at IPTG concentration 0.2 mM (lane 3), 0.4 mM (lane 5) and 0.6 mM (lane 7). Lane 1: Precision Plus protein standard (BioRad).

Appendix D2: Gene sequence for synthesis and expression in *E. coli*

Vector Map of pET16b (Novagen)

pET-16b sequence landmarks	
T7 promoter	466-482
T7 transcription start	465
His•Tag coding sequence	360-389
Multiple cloning sites	
(*Nde* I -*Bam*H I)	319-335
T7 terminator	213-259
lacI coding sequence	869-1948
pBR322 origin	3885
bla coding sequence	4646-5503



Original Gene Sequence before Modification

The part of the sequence highlited in red is the TAT signal sequence

>gb|CP002084.1|:1496492-1497940 Dehalogenimonas lykanthroporepellens BL-DC-9, complete genome

GAATTC

ATGAAATCGCATTCCACAATGAGTCGCCGAGATTTTATGAAGACCCTAGGTTTGGGTGCAACAGCCATTG
GTAGTGTAGGTGTCACCGCTCCCATTTTCCATGATCTCGACGAAATGATGAGCATAAGTGCTGCCGAAAC
TTTTAACTCCACTACATCCATGCAAAAACGTCCTTGGTGGGTTAAAGAGGTTGACATTCCTACAGTAGAA
ATCGACTTGAAACTACGTACCCCCTATGCGGGCCATCTATCTGCTGGTTTGTCTCCTCTATATCTTTCTA
AAGAGGAGATTGCTGCCATCCTGGCTTCTCAACAGAACAATGCTATTGAAGGAGCTAAAAACAACCGGCC
AGGTTTTACATTGCGTGATCAAATTGGAGCCTGGGCCTCGTTAGATAGAGGACAAACGGGATATGTGAAA
TATCCGCCTGAAGGTTTTCGAACAATAAAAGTTACTCATGAAACTTTAGGTGTACCGAAGTGGGAAGGCT
CCGAAACGGAAAACGCATTCATGATTCGAACTTTCCTGAGGCAATTTGGTGCGGGGGCGATTGGCTATGC
AAGAGTGGATGACGACAGCGTCGGACCACGTAAACCCCTTTTTAACACTCATGTAAGATTGGAGAACAAC
CCAGATTACAAGTATGACGCTAATGGTACATTCGTCATGCCAGAAAAGTGCAAGTATGCCATCATAATTT
ATGATAGAAGTCCCCGAGATCCTAACAACTATCGTCGTACCGTCAACAGCCCTCAAGCCTTTGTATCAAA
TATGGAGAAATGTGAGTATGGTCATAAGCTTCAAAACTTCCTTTGGGGTTTAGGCTACCAATCTTATTGG

TTTGAAGATGGCACAACAAGCAAATTCACCGGGACTCCGACGAATGTTTGGGGTATCCTCTCAGGCATAG
GCGAATACACCCGAATTCACAATCCGGTCTCGCAACCCGAAGGTGAGACCGGTAATTTTGCCAGTATTCT
CTTCACAGATTTGCCTCTACCTACTACTAAACCGATAGACTTTGGAGCTTTAGAATTCTGTAAAACTTGT
GGGATATGTGCCGACGTTTGCCCAGCCGGAGCAATTCCTACAGTAGAAGAGTATAGAGAGCCAACTTGGG
ATCGAGCAACTGGTCCCTGGAGTGCTTCCAATGACCATAAAGGATATCCTAATAAATCCATTGAATGCGT
AAAGTGGTATTTTTCTAATGCAGTTACAGCCTTCGCTCCTGCCTCTCGCCCAGTTGGTGTGTGTCGTCGA
TGTTCCAGCCATTGTGTCTTTAGTAAAGATCATAAAGCTTGGATTCATGAAGTAGTTAAGGGTGTAGTGT
CCACTACTCCTGTGATGAACAGCTTCTTTACTAAAATGGATACGCTATCCGGTTACAGTGACGTCATCTC
AGATGAAGGCAGGGCTGAATATTGGCACCAGTACCTGCCCGCTATTTAA
GGATCC

Optimized Sequence

(Optimized Sequence Length:1419, GC%:51.92)

CATATG
ACGCTGGGCCTGGGTGCCACGGCGATTGGTTCAGTTGGTGTTACGGCTCCGATCTTTCACGACCTGGACGAAATG
ATGAGTATCTCCGCGGCCGAAACCTTTAACTCAACCACGTCGATGCAGAAACGCCCGTGGTGGGTGAAAGAAGTT
GACATCCCGACGGTGGAAATTGATCTGAAACTGCGTACCCCGTATGCTGGCCATCTGAGTGCGGGTCTGTCCCCG
CTGTACCTGAGCAAAGAAGAAATCGCAGCTATTCTGGCGTCTCAGCAAAACAATGCCATCGAAGGCGCAAAAAAC
AATCGTCCGGGTTTTACCCTGCGTGACCAGATTGGTGCCTGGGCAAGCCTGGATCGTGGTCAAACGGGTTATGTG
AAATACCCGCCGGAAGGTTTTCGTACCATCAAAGTTACGCATGAAACCCTGGGCGTCCCGAAATGGGAAGGTTCA
GAAACGGAAAATGCTTTCATGATCCGTACCTTTCTGCGCCAGTTCGGCGCAGGTGCTATTGGCTATGCGCGTGTC
GATGACGATTCTGTGGGTCCGCGCAAACCGCTGTTTAACACGCACGTGCGTCTGGAAAACAATCCGGACTATAAA
TACGATGCCAATGGCACCTTCGTTATGCCGGAAAAATGTAAATACGCAATCATCATCTACGACCGTAGTCCGCGC
GATCCGAACAATTATCGTCGCACCGTGAACTCACCGCAGGCCTTTGTTTCGAATATGGAAAAATGCGAATACGGT
CATAAACTGCAGAACTTCCTGTGGGGCCTGGGTTATCAAAGTTACTGGTTTGAAGATGGCACCACGTCCAAATTC
ACCGGTACGCCGACCAATGTCTGGGGCATCCTGAGCGGCATTGGTGAATATACCCGCATTCACAACCCTGTTAGT
CAACCGGAAGGCGAAACGGGTAATTTTGCGTCCATCCTGTTCACCGACCTGCCGCTGCCGACCACGAAACCGATT
GATTTTGGTGCTCTGGAATTCTGCAAAACGTGTGGCATCTGCGCGGATGTTTGTCCGGCGGGTGCCATTCCGACC
GTCGAAGAATACCGTGAACCGACCTGGGACCGTGCAACCGGTCCGTGGAGTGCATCCAACGATCATAAAGGTTAC
CCGAACAAATCAATCGAATGTGTCAAATGGTACTTTTCGAATGCGGTGACCGCATTTGCTCCGGCCTCACGTCCG
GTTGGCGTCTGTCGTCGCTGCAGCTCTCACTGCGTGTTTTCGAAAGACCATAAAGCATGGATTCACGAAGTGGTT
AAAGGCGTCGTGAGCACCACGCCGGTTATGAACAGCTTTTTCACGAAAATGGATACCCTGAGCGGTTATTCTGAT
GTGATTAGCGATGAAGGTCGTGCGGAATACTGGCACCAATATCTGCCGGCGATTTGA
GGATCC

DNA Alignment (Optimized Region)

```
Optimized 7     ACGCTGGGCCTGGGTGCCACGGCGATTGGTTCAGTTGGTGTTACGGCTCCGATCTTTCAC
Original  7     ACCCTAGGTTTGGGTGCAACAGCCATTGGTAGTGTAGGTGTCACCGCTCCCATTTTCCAT

Optimized 67    GACCTGGACGAAATGATGAGTATCTCCGCGGCCGAAACCTTTAACTCAACCACGTCGATG
Original 67     GATCTCGACGAAATGATGAGCATAAGTGCTGCCGAAACTTTTAACTCCACTACATCCATG

Optimized 127   CAGAAACGCCCGTGGTGGGTGAAAGAAGTTGACATCCCGACGGTGGAAATTGATCTGAAA
Original 127    CAAAAACGTCCTTGGTGGGTTAAAGAGGTTGACATTCCTACAGTAGAAATCGACTTGAAA

Optimized 187   CTGCGTACCCCGTATGCTGGCCATCTGAGTGCGGGTCTGTCCCCGCTGTACCTGAGCAAA
Original 187    CTACGTACCCCCTATGCGGGCCATCTATCTGCTGGTTTGTCTCCTCTATATCTTTCTAAA

Optimized 247   GAAGAAATCGCAGCTATTCTGGCGTCTCAGCAAAACAATGCCATCGAAGGCGCAAAAAAC
Original 247    GAGGAGATTGCTGCCATCCTGGCTTCTCAACAGAACAATGCTATTGAAGGAGCTAAAAAC

Optimized 307   AATCGTCCGGGTTTTACCCTGCGTGACCAGATTGGTGCCTGGGCAAGCCTGGATCGTGGT
```

```
Original 307    AACCGGCCAGGTTTTACATTGCGTGATCAAATTGGAGCCTGGGCCTCGTTAGATAGAGGA

Optimized 367   CAAACGGGTTATGTGAAATACCCGCCGGAAGGTTTTCGTACCATCAAAGTTACGCATGAA
Original 367    CAAACGGGATATGTGAAATATCCGCCTGAAGGTTTTCGAACAATAAAAGTTACTCATGAA

Optimized 427   ACCCTGGGCGTCCCGAAATGGGAAGGTTCAGAAACGGAAAATGCTTTCATGATCCGTACC
Original 427    ACTTTAGGTGTACCGAAGTGGGAAGGCTCCGAAACGGAAAACGCATTCATGATTCGAACT

Optimized 487   TTTCTGCGCCAGTTCGGCGCAGGTGCTATTGGCTATGCGCGTGTCGATGACGATTCTGTG
Original 487    TTCCTGAGGCAATTTGGTGCGGGGGCGATTGGCTATGCAAGAGTGGATGACGACAGCGTC

Optimized 547   GGTCCGCGCAAACCGCTGTTTAACACGCACGTGCGTCTGGAAAACAATCCGGACTATAAA
Original 547    GGACCACGTAAACCCCTTTTTAACACTCATGTAAGATTGGAGAACAACCCAGATTACAAG

Optimized 607   TACGATGCCAATGGCACCTTCGTTATGCCGGAAAAATGTAAATACGCAATCATCATCTAC
Original 607    TATGACGCTAATGGTACATTCGTCATGCCAGAAAAGTGCAAGTATGCCATCATAATTTAT

Optimized 667   GACCGTAGTCCGCGCGATCCGAACAATTATCGTCGCACCGTGAACTCACCGCAGGCCTTT
Original 667    GATAGAAGTCCCCGAGATCCTAACAACTATCGTCGTACCGTCAACAGCCCTCAAGCCTTT

Optimized 727   GTTTCGAATATGGAAAAATGCGAATACGGTCATAAACTGCAGAACTTCCTGTGGGGCCTG
Original 727    GTATCAAATATGGAGAAATGTGAGTATGGTCATAAGCTTCAAAACTTCCTTTGGGGTTTA

Optimized 787   GGTTATCAAAGTTACTGGTTTGAAGATGGCACCACGTCCAAATTCACCGGTACGCCGACC
Original 787    GGCTACCAATCTTATTGGTTTGAAGATGGCACAACAAGCAAATTCACCGGGACTCCGACG

Optimized 847   AATGTCTGGGGCATCCTGAGCGGCATTGGTGAATATACCCGCATTCACAACCCTGTTAGT
Original 847    AATGTTTGGGGTATCCTCTCAGGCATAGGCGAATACACCCGAATTCACAATCCGGTCTCG

Optimized 907   CAACCGGAAGGCGAAACGGGTAATTTTGCGTCCATCCTGTTCACCGACCTGCCGCTGCCG
Original 907    CAACCCGAAGGTGAGACCGGTAATTTTGCCAGTATTCTCTTCACAGATTTGCCTCTACCT

Optimized 967   ACCACGAAACCGATTGATTTTGGTGCTCTGGAATTCTGCAAAACGTGTGGCATCTGCGCG
Original 967    ACTACTAAACCGATAGACTTTGGAGCTTTAGAATTCTGTAAAACTTGTGGGATATGTGCC

Optimized 1027  GATGTTTGTCCGGCGGGTGCCATTCCGACCGTCGAAGAATACCGTGAACCGACCTGGGAC
Original 1027   GACGTTTGCCCAGCCGGAGCAATTCCTACAGTAGAAGAGTATAGAGAGCCAACTTGGGAT

Optimized 1087  CGTGCAACCGGTCCGTGGAGTGCATCCAACGATCATAAAGGTTACCCGAACAAATCAATC
Original 1087   CGAGCAACTGGTCCCTGGAGTGCTTCCAATGACCATAAAGGATATCCTAATAAATCCATT

Optimized 1147  GAATGTGTCAAATGGTACTTTTCGAATGCGGTGACCGCATTTGCTCCGGCCTCACGTCCG
Original 1147   GAATGCGTAAAGTGGTATTTTTCTAATGCAGTTACAGCCTTCGCTCCTGCCTCTCGCCCA

Optimized 1207  GTTGGCGTCTGTCGTCGCTGCAGCTCTCACTGCGTGTTTTCGAAAGACCATAAAGCATGG
Original 1207   GTTGGTGTGTGTCGTCGATGTTCCAGCCATTGTGTCTTTAGTAAAGATCATAAAGCTTGG

Optimized 1267  ATTCACGAAGTGGTTAAAGGCGTCGTGAGCACCACGCCGGTTATGAACAGCTTTTTCACG
Original 1267   ATTCATGAAGTAGTTAAGGGTGTAGTGTCCACTACTCCTGTGATGAACAGCTTCTTTACT

Optimized 1327  AAAATGGATACCCTGAGCGGTTATTCTGATGTGATTAGCGATGAAGGTCGTGCGGAATAC
Original 1327   AAAATGGATACGCTATCCGGTTACAGTGACGTCATCTCAGATGAAGGCAGGGCTGAATAT

Optimized 1387  TGGCACCAATATCTGCCGGCGATTTGA
Original 1387   TGGCACCAGTACCTGCCCGCTATTTAA
```

Protein Alignment (Optimized Region)

```
Optimized 1   TLGLGATAIGSVGVTAPIFHDLDEMMSISAAETFNSTTSMQKRPWWVKEVDIPTVEIDLK
Original 1    TLGLGATAIGSVGVTAPIFHDLDEMMSISAAETFNSTTSMQKRPWWVKEVDIPTVEIDLK

Optimized 61  LRTPYAGHLSAGLSPLYLSKEEIAAILASQQNNAIEGAKNNRPGFTLRDQIGAWASLDRG
Original 61   LRTPYAGHLSAGLSPLYLSKEEIAAILASQQNNAIEGAKNNRPGFTLRDQIGAWASLDRG

Optimized 121 QTGYVKYPPEGFRTIKVTHETLGVPKWEGSETENAFMIRTFLRQFGAGAIGYARVDDDSV
Original 121  QTGYVKYPPEGFRTIKVTHETLGVPKWEGSETENAFMIRTFLRQFGAGAIGYARVDDDSV

Optimized 181 GPRKPLFNTHVRLENNPDYKYDANGTFVMPEKCKYAIIIYDRSPRDPNNYRRTVNSPQAF
Original 181  GPRKPLFNTHVRLENNPDYKYDANGTFVMPEKCKYAIIIYDRSPRDPNNYRRTVNSPQAF

Optimized 241 VSNMEKCEYGHKLQNFLWGLGYQSYWFEDGTTSKFTGTPTNVWGILSGIGEYTRIHNPVS
Original 241  VSNMEKCEYGHKLQNFLWGLGYQSYWFEDGTTSKFTGTPTNVWGILSGIGEYTRIHNPVS

Optimized 301 QPEGETGNFASILFTDLPLPTTKPIDFGALEFCKTCGICADVCPAGAIPTVEEYREPTWD
Original 301  QPEGETGNFASILFTDLPLPTTKPIDFGALEFCKTCGICADVCPAGAIPTVEEYREPTWD

Optimized 361 RATGPWSASNDHKGYPNKSIECVKWYFSNAVTAFAPASRPVGVCRRCSSHCVFSKDHKAW
Original 361  RATGPWSASNDHKGYPNKSIECVKWYFSNAVTAFAPASRPVGVCRRCSSHCVFSKDHKAW

Optimized 421 IHEVVKGVVSTTPVMNSFFTKMDTLSGYSDVISDEGRAEYWHQYLPAI*
Original 421  IHEVVKGVVSTTPVMNSFFTKMDTLSGYSDVISDEGRAEYWHQYLPAI*
```

Appendix E1: Gel images supporting results reported in chapter 4

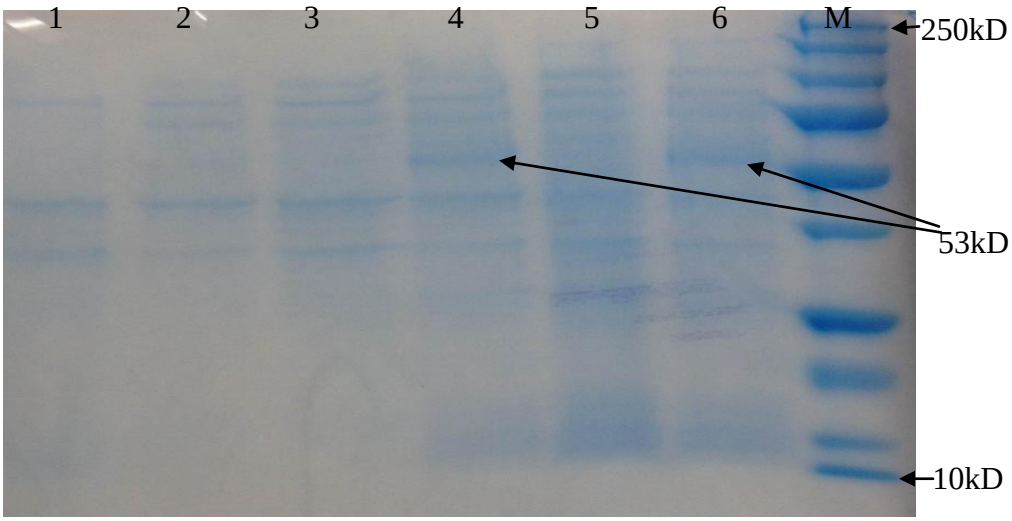


Fig E1. SDS PAGE analysis of the transformed *S. blattae* cells (co-transformed with Dehly_1524 and PceT). Lanes 4 and 6: cell lysate from anhydrotetracyclin induced *S. blattae* cells induced for 8 hours. Lanes 3 and 5: cell lysate generated from uninduced *S. blattae* cells grown under identical conditions. Lanes 1 and 2: soluble fraction of cell lysate recovered from anhydrotetracyclin induced *S. blattae* cells induced for 8 hours SDS-PAGE analysis was repeated to check for expression of the protein used for checking dehalogenase activity. Lane M: Precision Plus protein standard (BioRad).

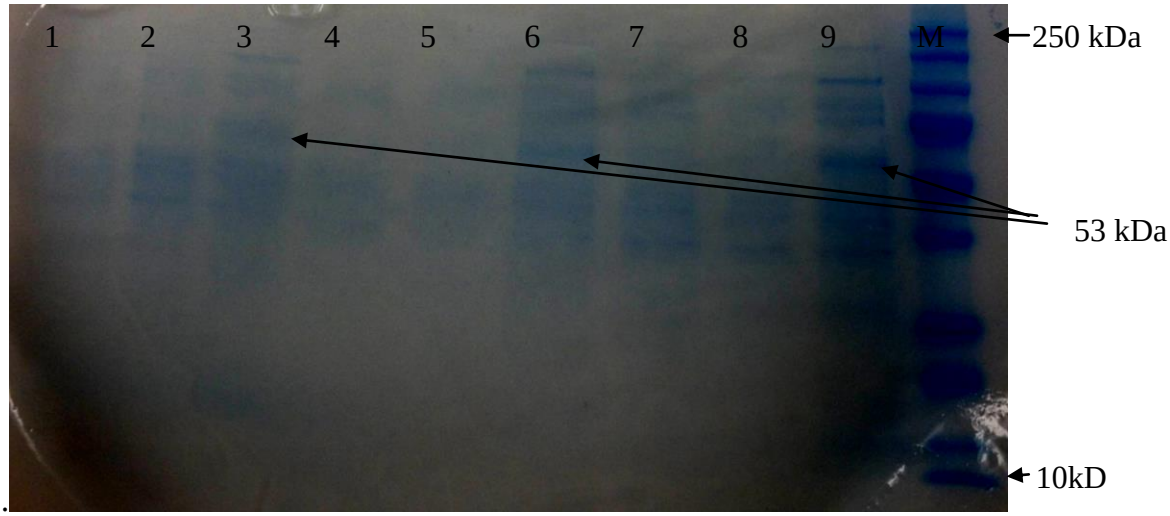
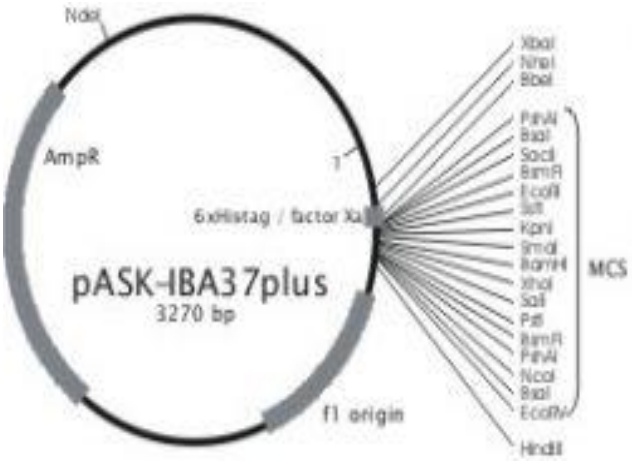


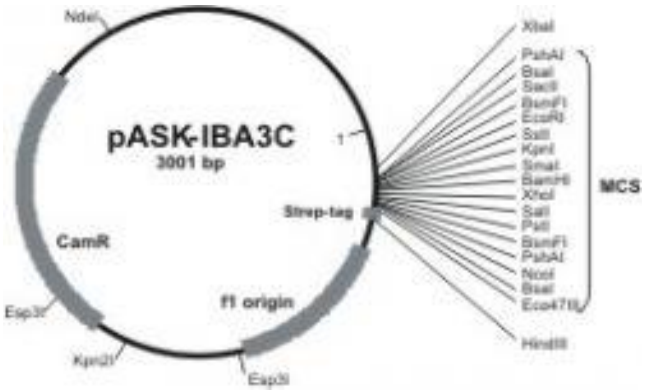
Fig E2. SDS PAGE analysis of the transformed *S. blattae* cells at lower induction concentration of 10 ng/mL anhydrotetracycline. Lanes 9, 6 and 3: cell lysate generated after induction for 7 Hours (lane 9), 5 hours (lane 6) and 3 Hours (lane 3). Lanes 1, 2, 4, 5, 7 and 8 are the soluble fractions of cell lysate recovered from cells induced for 3 Hours (lanes 1 and 2), 5 Hours (lanes 4 and 5) and 7 Hours (lanes 7 and 8). Lane M: Precision Plus protein standard (BioRad).

Appendix E2: Gene sequence for synthesis and expression in *S.blattae*

Vector Map of pASK-IBA 37 plus(IBA -Germany)



Vector Map of pASK-IBA 3C(IBA-Germany)



Original Gene Sequence before Modification

The part of the sequence highlighted in red is the TAT signal sequence

>gb|CP002084.1|:1496492-1497940 Dehalogenimonas lykanthroporepellens BL-DC-9, complete genome

ATGAAATCGCATTCCACAATGAGTCGCCGAGATTTTATGAAGACCCTAGGTTTGGGTGCAACAGCCATTGGTAGTGT
AGGTGTCACCGCTCCCATTTTCCATGATCTCGACGAAATGATGAGCATAAGTGCTGCCGAAAC
TTTTAACTCCACTACATCCATGCAAAAACGTCCTTGGTGGGTTAAAGAGGTTGACATTCCTACAGTAGAA
ATCGACTTGAAACTACGTACCCCCTATGCGGGCCATCTATCTGCTGGTTTGTCTCCTCTATATCTTTCTA
AAGAGGAGATTGCTGCCATCCTGGCTTCTCAACAGAACAATGCTATTGAAGGAGCTAAAAACAACCGGCC
AGGTTTTACATTGCGTGATCAAATTGGAGCCTGGGCCTCGTTAGATAGAGGACAAACGGGATATGTGAAA
TATCCGCCTGAAGGTTTTCGAACAATAAAAGTTACTCATGAAACTTTAGGTGTACCGAAGTGGGAAGGCT
CCGAAACGGAAAACGCATTCATGATTCGAACTTTCCTGAGGCAATTTGGTGCGGGGGCGATTGGCTATGC

AAGAGTGGATGACGACAGCGTCGGACCACGTAAACCCCTTTTTAACACTCATGTAAGATTGGAGAACAAC
CCAGATTACAAGTATGACGCTAATGGTACATTCGTCATGCCAGAAAAGTGCAAGTATGCCATCATAATTT
ATGATAGAAGTCCCCGAGATCCTAACAACTATCGTCGTACCGTCAACAGCCCTCAAGCCTTTGTATCAAA
TATGGAGAAATGTGAGTATGGTCATAAGCTTCAAAACTTCCTTTGGGGTTTAGGCTACCAATCTTATTGG
TTTGAAGATGGCACAACAAGCAAATTCACCGGGACTCCGACGAATGTTTGGGGTATCCTCTCAGGCATAG
GCGAATACACCCGAATTCACAATCCGGTCTCGCAACCCGAAGGTGAGACCGGTAATTTTGCCAGTATTCT
CTTCACAGATTTGCCTCTACCTACTACTAAACCGATAGACTTTGGAGCTTTAGAATTCTGTAAAACTTGT
GGGATATGTGCCGACGTTTGCCCAGCCGGAGCAATTCCTACAGTAGAAGAGTATAGAGAGCCAACTTGGG
ATCGAGCAACTGGTCCCTGGAGTGCTTCCAATGACCATAAAGGATATCCTAATAAATCCATTGAATGCGT
AAAGTGGTATTTTTCTAATGCAGTTACAGCCTTCGCTCCTGCCTCTCGCCCAGTTGGTGTGTGTCGTCGA
TGTTCCAGCCATTGTGTCTTTAGTAAAGATCATAAAGCTTGGATTCATGAAGTAGTTAAGGGTGTAGTGT
CCACTACTCCTGTGATGAACAGCTTCTTTACTAAAATGGATACGCTATCCGGTTACAGTGACGTCATCTC
AGATGAAGGCAGGGCTGAATATTGGCACCAGTACCTGCCCGCTATTTAA

Gene Sequence synthesized for subcloning along with the restriction sites

Sequence:

GAATTC
ATGAAATCGCATTCCACAATGAGTCGCCGAGATTTTATGAAGACCCTAGGTTTGGGTGCAACAGCCATTGGTAGTGT
AGGTGTCACCGCTCCCATTTTCCATGATCTCGACGAAATGATGAGCATAAGTGCTGCCGAAACTTTTAACTCCACTA
CATCCATGCAAAAACGTCCTTGGTGGGTTAAAGAGGTTGACATTCCTACAGTAGAAATCGACTTGAAACTACGTACC
CCCTATGCGGGCCATCTATCTGCTGGTTTGTCTCCTCTATATCTTTCTAAAGAGGAGATTGCTGCCATCCTGGCTTC
TCAACAGAACAATGCTATTGAAGGAGCTAAAAACAACCGGCCAGGTTTTACATTGCGTGATCAAATTGGAGCCTGGG
CCTCGTTAGATAGAGGACAAACGGGATATGTGAAATATCCGCCTGAAGGTTTTCGAACAATAAAAGTTACTCATGAA
ACTTTAGGTGTACCGAAGTGGGAAGGCTCCGAAACGGAAAACGCATTCATGATTCGAACTTTCCTGAGGCAATTTGG
TGCGGGGGCGATTGGCTATGCAAGAGTGGATGACGACAGCGTCGGACCACGTAAACCCCTTTTTAACACTCATGTAA
GATTGGAGAACAACCCAGATTACAAGTATGACGCTAATGGTACATTCGTCATGCCAGAAAAGTGCAAGTATGCCATC
ATAATTTATGATAGAAGTCCCCGAGATCCTAACAACTATCGTCGTACCGTCAACAGCCCTCAAGCCTTTGTATCAAA
TATGGAGAAATGTGAGTATGGTCATAAGCTTCAAAACTTCCTTTGGGGTTTAGGCTACCAATCTTATTGGTTTGAAG
ATGGCACAACAAGCAAATTCACCGGGACTCCGACGAATGTTTGGGGTATCCTCTCAGGCATAGGCGAATACACCCGA
ATTCACAATCCGGTCTCGCAACCCGAAGGTGAGACCGGTAATTTTGCCAGTATTCTCTTCACAGATTTGCCTCTACC
TACTACTAAACCGATAGACTTTGGAGCTTTAGAATTCTGTAAAACTTGTGGGATATGTGCCGACGTTTGCCCAGCCG
GAGCAATTCCTACAGTAGAAGAGTATAGAGAGCCAACTTGGGATCGAGCAACTGGTCCCTGGAGTGCTTCCAATGAC
CATAAAGGATATCCTAATAAATCCATTGAATGCGTAAAGTGGTATTTTTCTAATGCAGTTACAGCCTTCGCTCCTGC
CTCTCGCCCAGTTGGTGTGTGTCGTCGATGTTCCAGCCATTGTGTCTTTAGTAAAGATCATAAAGCTTGGATTCATG
AAGTAGTTAAGGGTGTAGTGTCCACTACTCCTGTGATGAACAGCTTCTTTACTAAAATGGATACGCTATCCGGTTAC
AGTGACGTCATCTCAGATGAAGGCAGGGCTGAATATTGGCACCAGTACCTGCCCGCTATTTAA
GGATCC

PceT gene from *D. hafniense* strain Y51

GAATTC

ATGATAATGAAGCAGTTCGAACTTGGACAATATAAGGGTCTAAATGTGAAGCGTTTTGATACAACAGTCCAAGAAGA
GGAGATTCAGCAAGCCTTAGATTATATCATAGGCTCTTTTGATGAAATTGAAGAAGAAAAAAGAAATGAGCCTATTA
AAACAAACGATTATGTAATCGTGGATATTGACGGTTATGAAAAGGATGCAACAGTACCGGTTATAAGAAACATCGAT
ACAAAATTAATAGTCGGCAGTGAAGGTGTATTTCGAGAAGTTAGTGCAAATCTCCTAGGGAAGAAAATGGGCGACAC
AGTGACCTTTGAGACTGTTATTCAGCCCGATGCCTTAGAATTTCAGCGTTGGTGGGGAAGCGAGTTTACATTTACCG
TTAAGATTAAGAGCGTTTTTGTGGTGAAGAAACCAGAATTGACTGAAGAATTAATTAGGAAAGTTGAACCTGATCTA
AAAAACCTGAAAGACTTAAAAAATATGCTTGCCCTAAAAATCACTCACGAAAAAGAGGGGAAGGAACGGGAGGCAAA
TATTCTCTTGATTTTTCAAGCTTTAGTTAAACAATGTAAGTATGAATTTGATGAAGAAGAATTGGATTCAGCTGCTG
AAGATTTATATAAGAAATTTACTGAAGAACTCAAAATAGTTGATGATATGGAACTTATGGAATATTTGATCCATAGA
AAAATAACCGCAGACCAATTGCTGGCCGAATGTAAGGAAGAGGCATCCAGAAGGATTCTCTGGGAACTTATGATAAA
TTCTGTTATTGAGAAAGAAGAGATAAATCTTACTCCGGATGAAATCAAATATCTTGAAAAGCGAATTAACGAAAGCC
GACAAAACGGTCAACTCCCCGAAGAGTTTATGGACATCAATTTTTTACTGGCATCGTATTTACGAAAAAAAACAATA
GACTATTTGCTAAAAATCAATTTGGCAAGTTAG

GGATCC

Permission from Annual Review of Biochemistry

Title: THE MANY FACES OF VITAMIN B12: CATALYSIS BY COBALAMIN-DEPENDENT ENZYMES 1

Author: Ruma Banerjee, Stephen W. Ragsdale

Publication: Annual Review of Biochemistry

Publisher: Annual Reviews

Date: Jun 1, 2003

Permission Not Required

Material may be republished in a thesis / dissertation without obtaining additional permission from Annual Reviews, providing that the author and the original source of publication are fully acknowledged.

Permission from Nature Publishing Group

NATURE PUBLISHING GROUP LICENSE TERMS AND CONDITIONS

Apr 17, 2015

This is a License Agreement between Kalpataru Mukherjee ("You") and Nature Publishing Group ("Nature Publishing Group") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3611390481718
License date	Apr 17, 2015
Licensed content publisher	Nature Publishing Group
Licensed content publication	Nature
Licensed content title	Reductive dehalogenase structure suggests a mechanism for B12-dependent dehalogenation
Licensed content author	Karl A. P. Payne, Carolina P. Quezada, Karl Fisher, Mark S. Dunstan, Fraser A. Collins, Hanno Sjuts
Licensed content date	Oct 19, 2014
Volume number	517
Issue number	7535
Type of Use	reuse in a dissertation / thesis
Requestor type	academic/educational
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	3
High-res required	no
Figures	Figure 2/ Crystal structure of NpRdhA in the resting state

Author of this NPG article	no
Your reference number	None
Title of your thesis / dissertation	MECHANISTIC INSIGHTS INTO CHLORINATED ALIPHATIC ALKANE REDUCTIVE DEHALOGENATION BY DEHALOGENIMONAS LYKANTHROPOREPELLENS BL-DC-9T AND DEVELOPMENT OF NUCLEIC ACID-BASED TECHNIQUES FOR MONITORING GENE EXPRESSION
Expected completion date	Aug 2015
Estimated size (number of pages)	120
Total	0.00 USD

Terms and Conditions

Terms and Conditions for Permissions

Nature Publishing Group hereby grants you a non-exclusive license to reproduce this material for this purpose, and for no other use,subject to the conditions below:

1. NPG warrants that it has, to the best of its knowledge, the rights to license reuse of this material. However, you should ensure that the material you are requesting is original to Nature Publishing Group and does not carry the copyright of another entity (as credited in the published version). If the credit line on any part of the material you have requested indicates that it was reprinted or adapted by NPG with permission from another source, then you should also seek permission from that source to reuse the material.

2. Permission granted free of charge for material in print is also usually granted for any electronic version of that work, provided that the material is incidental to the work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version.Where print permission has been granted for a fee, separate permission must be obtained for any additional, electronic re-use (unless, as in the case of a full paper, this has already been accounted for during your initial request in the calculation of a print run).NB: In all cases, web-based use of full-text articles must be authorized separately through the 'Use on a Web Site' option when requesting permission.

3. Permission granted for a first edition does not apply to second and subsequent editions and for editions in other languages (except for signatories to the STM Permissions Guidelines, or where the first edition permission was granted for free).

4. Nature Publishing Group's permission must be acknowledged next to the figure,

table or abstract in print. In electronic form, this acknowledgement must be visible at the same time as the figure/table/abstract, and must be hyperlinked to the journal's homepage.

5. The credit line should read:
   Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication)
   For AOP papers, the credit line should read:
   Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME], advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].XXXXX)

   Note: For republication from the *British Journal of Cancer*, the following credit lines apply.
   Reprinted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)For AOP papers, the credit line should read:
   Reprinted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME], advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].XXXXX)

6. Adaptations of single figures do not require NPG approval. However, the adaptation should be credited as follows:

   Adapted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication)

Trace element solution composition

HCL(25%=7.7 M	12.5 ml (100 mM)
$FeSO_4.7H_2O$	2100 mg (7.5 mM)
H_3BO_3	30 mg (0.5 mM)
$MnCl_2.4H_2O$	100 mg (0.5 mM)
$CoCl_2.6H_2O$	190 mg (0.8 mM)
$NiCl_2.6H_2O$	24 mg (0.1 mM)
$CuCl_2.2H_2O$	2 mg (0.01 mM)
$ZnSO_4.7H_2O$	144 mg (0.5 mM)

$Na_2MoO_4.2H_2O$	36 mg (0.15 mM)
Distilled Water	Fill to 1000 ml

Vita

Kalpataru Mukherjee is a native of the state of West Bengal from the Republic of India. He completed his schooling from St. Patricks Higher Secondary school, Asansol. Kalpataru Mukherjee received his Bachelor of Science in Biotechnology degree from Bangalore University, India in 2005 and his Master of Science in Biomedical Genetics degree from VIT University, India in 2007. He worked as a research assistant from December 2007 to April 2009 at Institute of Microbial Technology at Chandigarh, India and then enrolled in the doctoral programme in the Department of Biological Sciences at Louisiana State University and Agricultural and Mechanical College. Kalpataru Mukherjee is a candidate for the Doctor of Philosophy degree in Biological Sciences and expected to be awarded degree in August 2015.